UNIVERSITY OF READING

School of Agriculture, Policy, and Development

The Relationships between Genotype, Yield, and Environmental Factors on Protein Distribution in the Wheat Endosperm

by

George Peter Savill

Thesis submitted for the degree of Doctor of Philosophy

September 2018

Try to imagine a life without timekeeping.

You probably can't. You know the month, the year, the day of the week. There is a clock on your wall or the dashboard of your car. You have a schedule, a calendar, a time for dinner or a movie.

Yet all around you, timekeeping is ignored. Birds are not late. A dog does not check its watch. Deer do not fret over passing birthdays.

Man alone measures time.

Man alone chimes the hour.

And, because of this, man alone suffers a paralyzing fear that no other creature endures.

A fear of time running out.

Mitch Albom – *The Time Keeper*

Abstract

Distinct gradients exist in the distribution of protein within the wheat endosperm (Cobb 1905; Kent 1966; Tosi *et al.* 2011), with protein accumulation concentrated towards the outer endosperm. Due to these gradients, during milling mill streams of differing protein content and bread-making quality are produced (Wang *et al.* 2007). Furthermore, during milling for white flour a portion of the outer endosperm remains adhered to the removed aleurone layer, resulting in a disproportionate loss of protein.

By 2080, average summer temperatures in the UK are predicted to rise by up to 4.2°C, and heatwaves will become increasingly common (Jenkins *et al.* 2009). The effect that these changes will have on the distribution of protein within the wheat grain is yet to be discovered.

Controlled-environment and field trial experiments were used to investigate the effects of nitrogen input, temperature during grain-filling, and genotype on wheat grain protein distribution, protein quality and related gene expression, grain yield, and yield components. In addition to established techniques, a novel image analysis method was developed to quantify the protein distribution gradients in light-microscopy images of wheat grain.

Across both field and controlled-environment experiments, increasing nitrogen supply during vegetative development resulted in an increase in the protein distribution gradient in the wheat endosperm, with a greater proportion of protein accumulated towards the aleurone layer. Furthermore, elevated temperature during grain-filling was found to interact with nitrogen input, increasing the gradient in protein accumulation to a greater degree under high-nitrogen input conditions. In the field, the response to nitrogen varied between genotypes, with the high-protein bread-making varieties showing the greatest response. These grain protein distribution gradient results are supplemented by analyses of the size-distribution of protein bodies in the wheat endosperm, grain protein composition with regards to bread-making quality, gluten protein synthesis gene expression, and grain yield and yield components.

Acknowledgements

First and foremost I wish to thank my academic supervisors: Malcolm Hawkesford, for his excellent guidance and insight, continuous monitoring and support, providing me with uninhibited access to resources, and for allowing me to be part of his amazing team of scientists; Paola Tosi, for her thorough feedback on my written work, for sharing her expertise in microscopy and protein analysis, and for always being but a phone call away; and Peter Buchner, for his seemingly infinite knowledge of plant biochemistry, his patience and teaching ability in the laboratory, and the countless laughs he has provided over the past four years.

My thanks also go to Donna Fellowes, for giving me such a warm welcome on my first day, and for continued support throughout my studentship; and to Steve Thomas, for being a friendly and approachable mentor, for providing morale boosts when necessary, and for entrusting me with cooking duties at the annual departmental barbecue.

For their considerable contributions to my work whilst at Rothamsted Research, I would like to thank Adam Michalski, for developing the image analysis software which was critical to the success of this project; Andrew Riche, for ensuring the smooth running of the WGIN diversity field experiment; Anne Roßmann, for the incredible amount of work she put in to the collecting and processing of field samples during her visit in 2016, and for her continued intellectual input into my work; Chris Hall, for his extensive knowledge of cereal grain processing and bicycle maintenance; Kirstie Halsey, for teaching me everything I know about microscopy; Saroj Palmer, for both conducting all of my grain nitrogen content analyses, and for providing me with many delicious lunches; Stephen Powers, for his numerous in-depth meetings discussing the design and analysis of my experiments; Yongfang Wan, for training me in SDS-PAGE protein analysis, and for her eagerness to help my research in any way she can; Zhiqiang Shi, for guiding me through the process of SE-HPLC analysis and for his incredible enthusiasm for helping others; and to everyone else at Rothamsted that made my time there so pleasant and enjoyable.

I am also grateful to Maryna Kuzmenko and Andrii Seleznov, for inviting me to complete a thoroughly enjoyable internship at their software company, Petiole Ltd.

And finally, thank you to Lucas Gent, Nick Evens, and Yan Ma, for providing me with daily doses of distraction.

Strategic funding was provided by the Biotechnological and Biological Sciences Research Council (BBSRC) as part of the 20:20 Wheat (BBS/E/C/00005202) and Designing Future Wheat (BBS/E/C/000I0220) projects, the Lawes Trust, and the University of Reading.

Declaration of original authorship

Declaration: I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

George Savill

Table of Contents

List of Figures

List of Tables

Glossary

AHDB Agriculture and Horticulture Development Board.

ANOVA analysis of variance.

cDNA complementary DNA.

CSV comma separated values.

CV canonical variate.

DEPC diethylpyrocarbonate.

DF degrees of freedom.

dH_2O distilled water.

DNA deoxyribonucleic acid.

dNTP deoxynucleotide.

DPA days post anthesis.

DTT dithiothreitol.

GDH growing degree hours.

GS growth stage.

HGCA Home Grown Cereals Authority.

HMW high molecular weight.

HMW-GS high molecular weight glutenin subunit.

IAA isoamyl alcohol.

IPCC Intergovernmental Panel on Climate Change.

LMW low molecular weight.

LMW-GS low molecular weight glutenin subunit.

LSD least significant difference.

mRNA messenger ribonucleic acid.

NABIM National Association of British and Irish Flour Millers.

NRQ normalised relative quantification.

NUE nitrogen use efficiency.

PCR polymerase chain reaction.

qPCR quantitative PCR.

REML restricted maximum likelihood.

RNA ribonucleic acid.

RO reverse osmosis.

rRNA ribosomal RNA.

RT room temperature.

SDS sodium-dodecyl-sulphate.

SDS-PAGE sodium-dodecyl-sulphate polyacrylamide gel electrophoresis.

SED standard error of the difference.

SE-HPLC size-exclusion high-performance liquid chromatography.

SEM standard error of the mean.

TAE tris-acetate-EDTA.

TCA trichloroacetic acid.

TFA trifluoroacetic acid.

TGW thousand grain weight.

UV ultra-violet.

WGIN Wheat Genetic Improvement Network.

Chapter 1: Introduction

Earth's climate is changing. The temperature of the atmosphere and oceans is increasing, and the recent aim of limiting global warming to 1.5°C above pre-industrial levels, as specified by the Intergovernmental Panel on Climate Change (IPCC), will likely fail (Brown *et al.* 2017; Mauritsen *et al.* 2017; Raftery *et al.* 2017). Whilst the effects of climate change will be varied and far-reaching, perhaps the greatest effects will be on agriculture and food production. Current predictions show that crop yields will decrease (Parry *et al.* 2004; Lobell *et al.* 2007; Rosenzweig *et al.* 2014), driven not only by physiological responses to the environment, but also by an increase in the prevalence of pests and pathogens (Bebber *et al.* 2013).

Although climate change is the biggest challenge facing agriculture in the future, it is not the only challenge. The human population is increasing exponentially, and if this trend is to continue, food production must increase with it. However, this comes at a time when the impact of agriculture on the environment must be minimised through the reduction of energy-intensive inputs and improvements in land-use management. In short, agricultural productivity needs to continue to increase, but with fewer inputs, using the same amount of land, and whilst withstanding increasingly unfavourable environmental conditions.

1.1. An introduction to wheat

Wheat is the largest crop in the UK, the third biggest cereal crop worldwide (FAOSTAT 2014), and is grown in an unparalleled range of environments and climates (Shewry 2009). As a global crop, wheat represents a major source of dietary protein and carbohydrate, and its grain is used to make a vast range of foods. Its success as a food crop is due largely to the nutritional and processing qualities of its grain and flour. It is as flour, either wholegrain or processed, that the majority of wheat grain is consumed (Peña 2002). Wheat flour is high in protein compared to that of other cereals, and is defined by its large gluten protein faction which allows for the formation of a visco-elastic dough upon wetting with water (Lásztity 1996). The abundance of these gluten proteins facilitates the creation of a dough that is both elastic and extensible, and which allows for the production of a range of baked foods including leavened (fermented) bread, pasta, noodles, and biscuits (Carson *et al.* 2009). Both the total protein content and protein composition of the mature wheat grain vary between cultivars, and determine the suitability of the produced flour for different applications based on certain end-use requirements (Morris *et al.* 1996).

Modern wheat varieties belong to one of two species, the hexaploid common wheat (*Triticum aestivum* L.) or tetraploid durum wheat (*Triticum durum* L.). In addition to these taxonomic groups, wheat can be categorised by their protein content and 'hardness' of the grain, which ultimately determine the end-use of the grain produced. Durum, or pasta wheats have extremely hard grains, with high gluten and total protein content. They account for approximately 5%

of wheat grown worldwide, and are used primarily to produce pasta food products (Morris *et al.* 1996; Shewry 2009). Common, or bread wheat makes up the remaining 95% of wheat grown, and is used to produce a much wider range of food products including leavened bread. This study will focus solely on *Triticum aestivum*.

In the UK, common wheat varieties are categorised into groups in accordance with their end-uses, as specified by National Association of British and Irish Flour Millers (NABIM) (NABIM 2014). The NABIM rating system takes into account protein content and grain texture to place each variety of wheat into one of four groups. Group one and two wheat varieties are hard bread-making wheats with high protein content: group one consist of the best bread-making varieties, which offer consistently high performance, whilst group two wheats are less consistent, either with undesirable quality traits, or else they only produce flour of bread-making quality under the most favourable conditions. Hence, group two wheats generally command lower prices than group one wheats. Group three wheat varieties are soft-grained, with lower protein content, and are suited to production of cakes and biscuits as well as for use in distilling. Finally, group four wheats can sub-categorised as either hard or soft-grained, have the lowest protein content, and are generally used for animal feed. The NABIM rating system can be summarised by saying wheat varieties of a lower group number are more valuable and have a higher grain protein content.

1.2. **Wheat morphology**

The wheat grain, or caryopsis, is a single-seeded fruit contained within the ear of the wheat plant. Each ear contains multiple grain, which contain the large central starchy endosperm and the embryo, encased in the protein-rich aleurone and bran layers which coat the entire grain. The endosperm acts as a storage organ for the grain, and is rich in the protein and carbohydrates that make wheat such a useful and nutritious crop. It is the physical structure and chemical composition of the grain, as determined by the process of grain development, that determine the yield, value and end-use suitability of the mature grain.

1.2.1. Morphology of the ear

The wheat inflorescence, known as the ear or spike, is the organ which contains the individual flowers, which become developing grain after fertilisation. The spike consists of two rows of alternate spikelets connected by the central rachis. Each spikelet contains up to six florets which each produce a single mature grain (Shewry *et al.* 2009a). Figure 1.1 shows a graphical representation of a mature ear and spikelet. The size of the grain within each spikelet varies, with the florets towards the base maturing faster and producing larger grain. Likewise, maturation occurs at different rates within the ear, with central spikelets maturing faster than those at the base and tip. This behaviour is also reflected in the progression of anthesis, which is illustrated in figure 1.2.

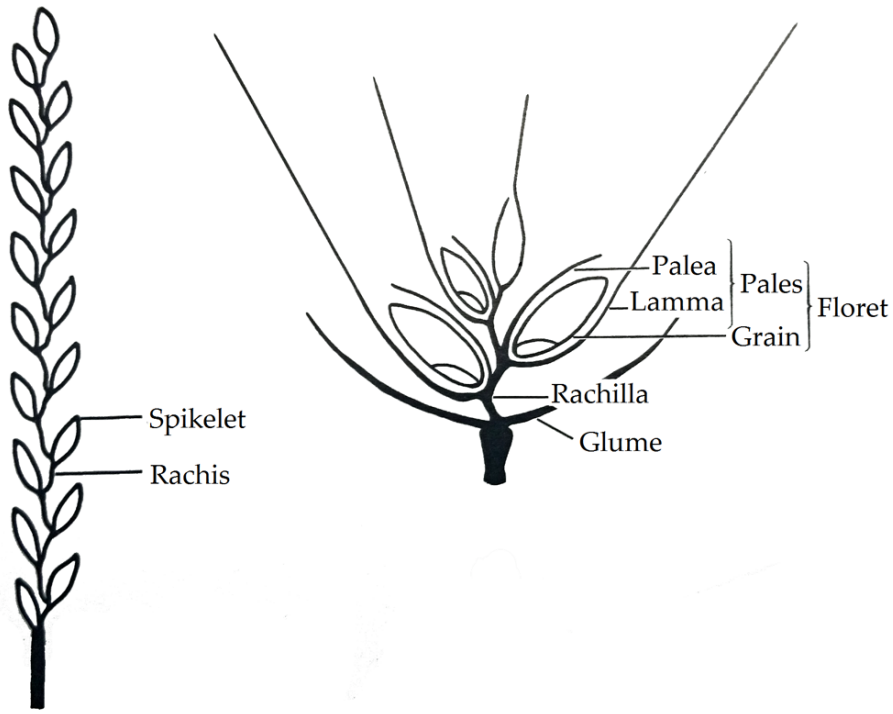


Figure 1.1: The mature wheat ear (left) and spikelet (right). Each ear contains multiple spikelets supported on a central rachis. Each spikelet contains multiple grain protected by the palea and lamma. Adapted from **Belderok2000**.

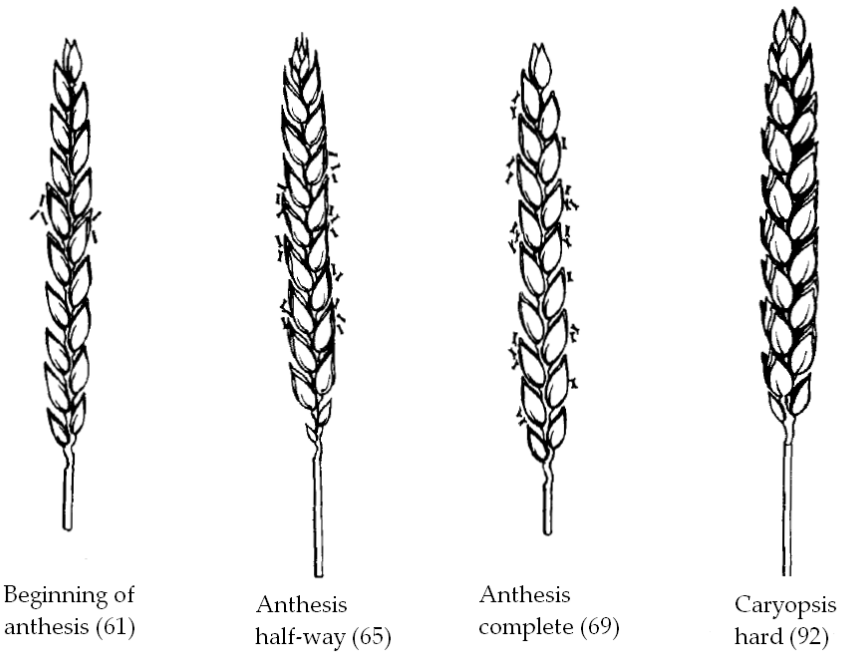


Figure 1.2: The progression of anthesis from Zadok's growth stage 61 to 69, and the mature harvest-ripe ear at growth stage 92. Anthesis begins in the central spikelets of the wheat ear, and spreads to the ends of the ear over the course of several days. Adapted from Tottman 1987.

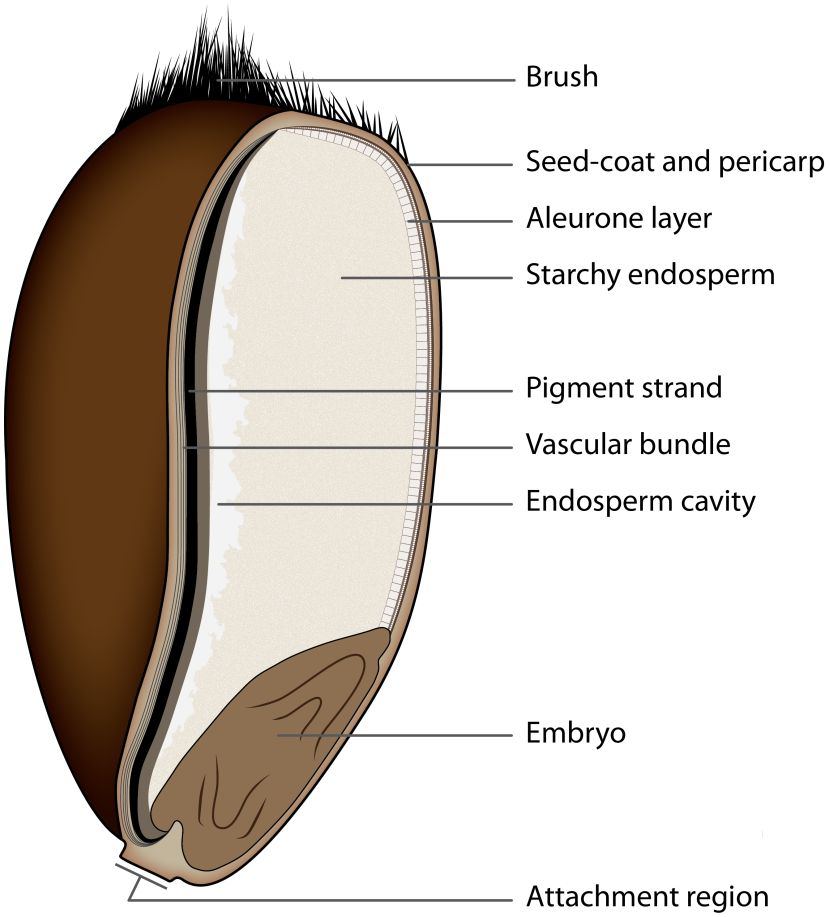


Figure 1.3: A cross-section of the mature wheat grain showing the internal structure. At the base of the grain is the embryo and the attachment region. Inside the crease, running the length of the grain, is the vascular bundle, pigment strand, and endosperm cavity. The bulk of the grain consists of the starchy endosperm, which is encased in the aleurone, seed-coat, and pericarp layers.

1.2.2. Morphology of the grain

The wheat grain is approximately oval, with a longitudinal crease running the length of the ventral side which reaches almost into the centre of the grain. The embryo is located on the dorsal side at the basal end, next to the attachment region (a scar formed from where the developing grain was attached to the wheat plant). At the apical end, opposite from the attachment region, is a tuft of hairs which constitute the brush. The gross morphology of the grain varies among different wheat cultivars, with some varieties showing more elongated grain. Likewise, when observed in transverse cross-section, grain can take on shapes from almost triangular to circular. However, the general characteristics described here are common across all varieties.

The interior of the grain consists of the endosperm, which contains a central storage body known as the starchy endosperm which is rich in starch and protein, with clear gradients between these two components (Kent *et al.* 1969; Tosi *et al.* 2011; He *et al.* 2013). This

starchy endosperm is coated with a single layer of protein rich cells called the aleurone layer. The endosperm tissues are encased in the pericarp, formed by the ripened ovary wall and consisting of an outer epidermis, hypodermis, parenchyma, intermediate cells, cross cells, and tube cells (Shewry *et al.* 2009a). The pericarp is relatively low in protein, and is removed during the production of flour. Figure 1.3 shows an illustration of a mature wheat grain in lateral cross-section.

1.3. Development of the wheat plant

The quality and final protein content of the wheat grain is a result of the growth and development of the wheat plant, from germination through to harvest: it is the photosynthate accumulated, and nutrients absorbed during this period which ultimately form the grain post-fertilisation.

Prior to the development of grain, wheat goes though several stages of vegetative development. Following germination, seedlings emerge and possess a single tiller. During early development 'tillering' occurs where multiple tillers are formed and continue to grow. At this stage winter wheat varieties overwinter and vernalise, whereas spring wheats—which do not require vernalisation—continue straight into the developmental stage of stem elongation. The stems elongate rapidly as temperatures increase before producing a flag leaf. Once a flag leaf is established the inflorescence, or ear, begins to form within during a process known as booting. During booting the ear develop within the protective sheath of the flag leave prior to emergence and subsequent anthesis (or flowering). Anthesis quickly leads to self-pollination which initiates grain development (discussed in section 1.3.1).

The development of wheat can be described and quantified using the Zadoks scale for cereal growth stages (Zadoks *et al.* 1974), which is a scale developed specifically for the measurement of easily recognisable developmental stages to provide detailed information which can be used for analysis of cereal development. The ten major groups or stages of the scale are show in table 1.1, with table 1.2 showing the more detailed individual scores from anthesis through to harvest ripeness. Zadok's growth stages 60 to 92 are of interest to this study, since they cover the period from anthesis to harvest ripeness.

1.3.1. Grain development

The development of the wheat grain takes place in the time between fertilisation and harvest, where the grain is at full harvest maturity (Zadok's GS 92). The process of grain development can be divided into three distinct biological stages: the first stage of cell division, the second stage of cell enlargement and the final stage of dehydration and grain maturation (Carceller *et al.* 1999). It is during this time that the starch- and protein-rich endosperm is formed, creating the nutritionally valuable final product that is the mature wheat grain. Figure 1.4 shows a graphical representation of seed development during this period. Whilst the key events are common to all wheat varieties, the timings and rates of each stage of grain development vary between genotypes, and are greatly affected by environmental conditions (Sofield *et al.*

Table 1.1: Zadok's decimal scale for the growth stages of cereals from germination to seed dormancy. Adapted from Zadoks *et al.* 1974.

Growth stage		Description
$0n$	Germination	n = stages from dry seed to first leaf emerging from coleoptile
$1n$	Seedling growth	n= number of leaves unfolded
$2n$	Tillering	n= number of tillers
$3n$	Stem Elongation	n= number of nodes to a total of 6. 7-9 represent flag leaf emergence
$4n$	Booting	n = degree of swelling
$5n$	Ear emergence	n = level of inflorescence emergance
$6n$	Anthesis	n = stage of anthesis
$7n$	Milk development	Grain development from watery (n=0) to milky (n=7)
$8n$	Dough development	Grain development from early dough (n=3) to hard dough (n=7)
$9n$	Ripening	n=2 is harvest-ripe, n=4 is over-ripe and n=5 is seed dormancy

Table 1.2: Zadok's decimal scale for the growth stages of cereals during grain-filling. Growth stages are presented for anthesis, and early (milk development), mid (dough development), and late (ripening) grain-filling. Adapted from Zadoks *et al.* 1974 and Tottman 1987.

Growth stage	Description
Anthesis	
60	Beginning of anthesis
65	Anthesis half way
69	Anthesis complete
Milk development	
71	Caryopsis water ripe, grain is up to 3mm and contains colourless liquid
73	Early milk, grain contains white, watery liquid
75	Medium milk, grain nearly full length and contain a soft liquid centre
77	Late milk, grain contents are wet and sticky when crushed
Dough development	
83	Early dough, grain contents are soft and cheesy
85	Soft dough, grain contents are firm and difficult to squeeze out. Thumb-nail impression quickly disappears. Green colour begins to fade
87	Hard dough, grain contents are dry and impossible to squeeze out. Thumb-nail leaves impression. Green colour lost
Ripening	
91	Caryopsis hard, grain is difficult to divide with thumb-nail
92	Caryopsis hard, grain cannot be dented with thumb-nail. Grain is harvest-ripe
93	Caryopsis loosens during the day
94	Over-ripe, straw dead
95	Seed dormant
96	Viable seed showing 50% germination
97	Primary dormancy lost
98	Secondary dormancy induced
99	Secondary dormancy lost

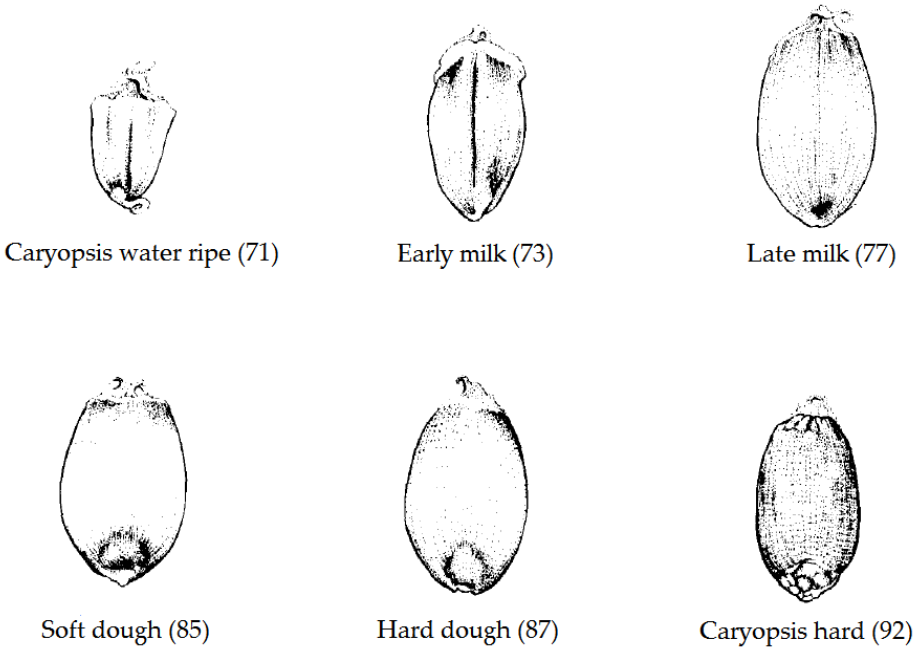


Figure 1.4: The external appearance of development wheat grain from early grain-filling to maturity. Stages of development from Zadok's growth stage (GS) 71 to 92 are shown (described in table 1.2), which cover the full range of grain development. Illustrated grain are shown from the dorsal side. Adapted from Tottman 1987.

1977).

1.3.1.1. *Cell division*

The initial stage of grain development is cell division, induced by the fusion of male and female gametes during fertilisation. The cell division phase of development is initiated just hours after fertilisation, lasts approximately 15–20 days, and begins with the division of the endosperm nucleus which will go on to form the endosperm of the grain. Rapid water uptake occurs in coordination with a period of intense mitotic cell division, resulting in a vast increase in grain dry matter, water content, and total cell number.

The increase in dry matter during this period of development is due largely to the production of starch and protein. Within the first few days, A-type starch granules begin to form, and reach their final size prior to the end of the cell division stage at around 19 DPA. B-type starch granules are initiated later, increasing in size from around 21 DPA through to the later stages of grain development (Bechtel *et al.* 1990; Jenner *et al.* 1991). In addition to starch, proteins also begin to accumulate within the developing grain. From 10 DPA onwards, amino acids are remobilised from proteins within the senescing leaf tissues, and storage proteins begin to form into discrete protein bodies within the developing grain (Jenner *et al.* 1991). During the later stages of grain development these protein bodies fuse to form a continuous protein matrix which embeds the starch granules.

By the end of this period of grain development mitotic cell division ceases, and a maximum cell number of approximately 10^5 is reached (Jenner *et al.* 1991; Gooding 2009). After this stage further grain growth will occur only via cell enlargement.

1.3.1.2. *Cell enlargement*

The cell enlargement phase of grain development overlaps with the cell division stage, beginning at around 10-15 DPA, and continuing for a further 15 to 30 days. The exact timing of this phase depends heavily on temperature (Gooding 2009).

During the cell enlargement stage, the maximum cell number is reached, and further grain filling occurs only through the enlargement of existing cells. During this period the water content of the developing grain is maintained at a relatively constant level, with little or no net water loss. Dry matter levels, however, continue to increase as starch and protein accumulation continues (Jenner *et al.* 1991; Pepler *et al.* 2006). The production of B-type starch granules continues, and smaller C-type granules begin to form at around 20 DPA (Bechtel *et al.* 1990). The protein bodies first observed during the initial stages of cell division continue to grow, and eventually fuse to form a protein matrix embedding the newly formed starch granules (Wang *et al.* 1995; Gooding 2009). This process results in a change in the consistency of the developing grain, going from a milky liquid to a more viscous dough-like texture.

1.3.1.3. *Dehydration and grain maturation*

The onset of rapid water loss from the developing grain marks the beginning of the dehydration and grain maturation phase of grain development (Pepler *et al.* 2006). From onset until maturity the water content of the grain decreases rapidly, whilst dry matter content remains relatively stable. This results in a slight decrease in the size and weight of the grain, but a change in consistency. During this time the grain hardens significantly from a soft dough-like consistency to the dry, hard grain observed at harvest, which corresponds to a score of 92 on Zadok's cereal growth scale. The final product of grain development is a hard, dried grain, enriched with starch and proteins which is suitable for storage or processing.

1.4. **Nitrogen**

Nitrogen is an essential element for all life on earth, and is the limiting factor of crop productivity in the majority of unmanaged soils (Robertson *et al.* 2009). As such, modern agriculture, and the seven-plus billion people it supports, currently relies on the application of over a hundred million tons of nitrogen fertiliser per year, with that figure increasing year-on-year (Tilman 1999; FAOSTAT 2014). In the UK, nitrogen fertiliser use is particularly high, and farmers currently apply 250-300kg-N/ha to produce a crop with the 13% protein content required for bread-making (Shewry 2009).

Both the production, and the application of such vast quantities of nitrogen fertiliser to our farmland poses a massive environmental cost (discussed further in section 1.8.1), and as such increasing the nitrogen use efficiency (NUE) of crop our crops should be a major priority

for the agriculture industry. Regardless of the moral issues surrounding the environmental issues surrounging nitrogen fertiliser use, it is likely that farmer's ability to use high levels of nitrogen fertilisers will be restricted in the future, either through increases in cost of product, or due to political efforts to limit the environmental impact of agriculture.

As discussed in section 1.1, the value of a wheat crop is largely determined by the protein content of its grain. Whilst the amount of nitrogen supplied to a wheat crop determines the crop's development, yield, and grain protein content (Hussain *et al.* 2006; Otteson *et al.* 2007), through the differential effect of nitrogen input on the accumulation of differen gluten proteins, it also determines the crop's suitability for use in bread-making (Godfrey *et al.* 2010; Chope *et al.* 2014). Therefore reducing the amount of nitrogen supplied to a wheat crop would be detrimental in terms of grain yield, nutritional quality, and in the value of the grain produced. As such it would be beneficial to further our understanding of the effect nitrogen supply has on the quality of wheat grain, to allow for the production of high-protein bread-making wheat that require lower levels of nitrogen input.

1.4.1. The transport of nitrogen from the rhizosphere to the grain

In wheat, nitrogen is absorbed as either nitrate or ammonium by the root system (where some is assimilated into the roots) before being transported to the shoot. Nitrate is reduced to nitrite in the cytoplasm, before further reduction to ammonium. Ammonium, either directly absorbed by the roots or else assimilated from nitrate, is then used in the production of the amino acids glutamine, glutamate, asparagine, and aspartate (Lam *et al.* 1996). These amino acids are then transported to the developing leaves, which act as a strong nitrogen sink during vegetative development (Okumoto *et al.* 2011).

Whilst the developing leaves act as a sink for amino acids during vegetative development, after anthesis the developing grain become a sink for nitrogen. Although a small amount of nitrogen is supplied directly from the roots in the form of nitrate and ammonium, the majority (up to 70%) is supplied by amino acids remobilised from elsewhere in the plant (Kichey *et al.* 2007; Pask *et al.* 2012). As grain-filling commences, senescence of leaf material begins, and proteins within the leaf tissue are degraded into amino acids which are then transported to the developing grain (Gregersen *et al.* 2008). Bancal (2009) observed that the amount of nitrogen remobilised into the developing grain is ultimately determined by the amount of nitrogen absorbed prior to anthesis, and so increasing the accumulation of nitrogen in the plant prior to anthesis would be beneficial in increasing grain protein content at harvest.

Whilst the process of nitrogen assimilation, transport, and remobilisation around the wheat plant is well understood, investigations into the mechanisms behind the production of distinct patterns in protein accumulation in the grain have only recently begun (Moore *et al.* 2016), the findings of which are discussed further in section 1.7.

1.5. Composition of the wheat grain

The versatility of the wheat crop is a result of the variable composition of its grain. As a product of grain development, the composition of the wheat grain is greatly affected by the nutrition and environmental conditions experienced during development (Hurkman *et al.* 2013; Wan *et al.* 2013; Wan *et al.* 2014). Likewise, different varieties produce grain with vastly different grain compositions (Bergman *et al.* 1998; Chope *et al.* 2014), from high-protein bread-making wheats to low-protein, high-starch feed varieties. The ultimate composition of the harvested grain dictates the end-use, and therefore value, of the crop.

1.5.1. Carbohydrates

Carbohydrates are the largest component of the grain, with mature wheat grain containing approximately 85% carbohydrates. Of this, ∼80% is starch, which is found solely in the endosperm of the grain, ∼7% is low molecular mass mono-, di-, and oligosaccharides, which are present in the endosperm, aleurone layer and embryo tissues, and ∼12% is cell wall polysaccharides, which are found in all tissues (Stone *et al.* 2009). Of these carbohydrates, it is the starch which is of most importance with regards to the nutritional and processing qualities of the grain.

1.5.1.1. *Starch*

In bread-making, the starch content of the mature grain contributes significantly to the baking characteristics of wheat flour. Starch provides the sugars required for yeast-fermentation, and dictates the gelatinisation characteristics of the flour. Hence, the starch content of the grain dictates the crumb texture and staling of bread (Yasunaga *et al.* 1968), and provides the structural strength required to make biscuits and cakes. Furthermore, starch constitutes a major source of dietary calories, with wheat providing approximately one-fifth of the calories consumed by humankind globally (Rasheed *et al.* 2014).

Biologically starch acts as the primary energy storage of the mature grain. Accumulated during grain development, starch is formed primarily from the sucrose produced by photosynthesis, but also from the remobilisation of other carbohydrates stored within the plant (Stone *et al.* 2009). In the early days of grain development, during the cell division stage, amyloplasts begin to form large A-type starch granules, which reach a maximum diameter of 25-50 µm from 19 days after anthesis. Smaller, B-type starch granules are initiated later, at approximately 10 DPA. These B-type granules begin to enlarge at around 21 days after anthesis, reaching a final size of on 9 µm in diameter at maturity (Bechtel *et al.* 1990). Later in development (from 21DPA onwards), smaller C-type granules grow to a diameter of less than 5.3 µm. Although these smaller B and C-type granules are much more abundant than the A-type granules, due to the size of the A-type granules, these contain the majority of the starch present in the grain (Evers 1973; Bechtel *et al.* 1990; Jenner *et al.* 1991).

Due to the order and nature of starch granule production, a starch gradient is established within the starchy endosperm. Since the large A-type starch granules, which account for the

majority of deposited starch, are formed early in development, they are concentrated within the older tissues of the grain, i.e. in the centre. As the grain develops and grows, the smaller B and C-type granules are produced in the newer tissues of the grain. Since they account for a smaller proportion of the total grain starch content, there is therefore less starch in the newer parts of the grain. Hence, a clear starch gradient is present in mature wheat grains, with a higher concentration of starch at the centre of the starchy endosperm, and lower concentrations nearer the aleurone layer (Jenner *et al.* 1991).

During milling, the wheat grain starch granules are damaged. The level of starch damage observed is correlated with the hardness of the grain. Softer wheats experience lower levels of starch damage than harder wheats, which in turn affects the processing properties of the obtained flour. Flour with higher levels of starch damage absorbs water more readily, and provides a greater source of fermentable sugars, which are required for the production of leavened bread (Carson *et al.* 2009). Conversely, flours from soft wheats contain lower levels of starch damage, absorb less water, and are more suited to the production of cakes and biscuits (Morris *et al.* 1996).

1.5.1.2. *Non-starch carbohydrates*

Although generally considered less significant nutritionally, the non-starch carbohydrates of the wheat grain confer various dough properties which impact on the end use of the grain. Such carbohydrates include the free sugars and non-starch polysaccharides within the endosperm and the cell walls of the grain (Henry 1985).

In high-protein bread-making wheat varieties, free sugars account for approximately 1-2% of the weight of the endosperm and act as a readily available substrate for the yeast during fermentation (D'appolonia *et al.* 1995). Additionally, other non-starch sugars such as cell wall arabinoxylans have also been shown to have an impact on baking quality (Courtin *et al.* 2002).

In low-protein wheats used in distilling for ethanol production, such non-starch carbohydrates have been shown to affect the alcohol yield achieved. Wheat cultivars with higher levels of non-starch carbohydrates show an inhibition to ethanol production and achieve a lower alcohol yield (Davis-Knight *et al.* 2007).

1.5.2. Protein

After carbohydrates the grain consists primarily of a range of proteins: metabolic, structural and storage proteins. The storage proteins are not only the most abundant group of proteins within the mature wheat grain, but are also the most important with regards to grain quality and end-use. For a grain protein to be classified as a storage protein it must meet the following criteria, as specified by Kreis *et al.* (1985):

- Tissue specificity: Protein is specific to the grain tissues

- Temporal accumulation: Protein accumulates during the later stages of grain development.
- Disproportionate response to nitrogen nutrition: Accumulation is disproportionally decreased or increased by limiting or abundant nitrogen supply respectively.
- Subcellular location: Protein accumulates to form discrete protein bodies within the cells of the grain.
- Absence of other function: Protein lacks any function other than as a store of nutrients.

Storage proteins begin to form during the early stages of grain development, and first appear at around 10 DPA in spherical membrane-bound vesicles, or protein bodies (Jenner *et al.* 1991). These vesicles are derived from the golgi apparatus and are closely associated with the rough endoplasmic reticulum (Miflin *et al.* 1981). By maturity, these distinct protein bodies are virtually absent from the grain (Payne *et al.* 1986), as the vesicles fuse to form larger, irregular protein bodies, which eventually fuse to form a continuous protein matrix which embeds the starch granules described in section 1.5.1.1 (Jenner *et al.* 1991; Shewry *et al.* 2009b). At maturity, the storage protein faction of the grain constitutes between 8% and 20% of the total dry matter. Both the final concentration and composition of protein depends heavily on genotype, nutrition, and the environmental conditions experienced during development (Shewry *et al.* 2009b).

Of the total grain protein fraction, approximately 80% to 85% are gluten storage proteins (Peña 2002). It is this group of proteins that confer the visco-elastic properties of the dough formed upon mixing wheat grain flour and water. Grain with a higher gluten protein content shows greater strength, extensibility, and elasticity, and is more suited to bread-making (El Haddad *et al.* 1995; Sapirstein *et al.* 1998). Grain with lower gluten protein content, however, is more suited to the production of cakes, biscuits, or animal feed. Therefore, the proportion of gluten protein within the endosperm of the grain relates directly to the final quality and hence the value of the mature grain product.

The gluten proteins which accumulate within the wheat grain can be categorised as either monomeric gliadins or polymeric glutenins (Shewry *et al.* 2009b). Within these two groups, the gluten proteins can be further sub-divided into several families. The glutenin proteins are categorised according to their molecular weights as either high molecular weight (HMW) or low molecular weight (LMW) subunits. Likewise, the gliadins are also further categorised as alpha-, beta-, gamma, or omega-gliadins (Kreis *et al.* 1985). It is both the quantity, and the ratio between these groups of storage proteins which dictate the processing qualities of the mature wheat grain (Khatkar *et al.* 1995; Uthayakumaran *et al.* 2000).

1.5.2.1. *Glutenins*

The glutenin protein polymers are elastic in nature, and confer strength and extensibility to the dough. This elasticity is due largely to the polymeric nature of the glutenin proteins: glutenin

polymers consist of high- and low-molecular-weight subunits that are linked by disulphide bonds to form molecules of expansive molecular weights. It is the quantity, composition and ratio of these two categories of subunits which interact to determine dough strength and extensibility (Khatkar *et al.* 1995; Zhang *et al.* 2007b; Zhang *et al.* 2007a).

High molecular weight glutenin subunits (HMW-GSs) account for approximately 12% of the total storage protein within the mature wheat grain, and are most important for dough strength (Rasheed *et al.* 2014). Encoded by genes at the *Glu-1* loci on the long arms of the homologous group 1 chromosomes named *Glu-A1*, *Glu-B1* and *Glu-D1* respectively. Each of these loci encode for the production of x and y-type subunits, with slightly higher and lower molecular weights respectively. Common to all bread-making wheat cultivars are the 1Bx, 1Dx and 1Dy subunits, whilst some bread-making wheats additionally contain the 1By, 1Ax and rare 1Ay subunits (Margiotta *et al.* 1996; Rasheed *et al.* 2014). Further to the different subunits produced by genes at different loci, these is also allelic variation within these loci, with the *Glu-A1*, *Glu-B1* and *Glu-D1* loci having 21, 69 and 29 known alleles respectively (Rasheed *et al.* 2014). Through this allelic variation there is scope for considerable variation in bread-making quality.

Low molecular weight glutenin subunits (LMW-GSs) represent approximately a third of the total storage proteins, make up around 60% of the glutenin protein group (Rasheed *et al.* 2014), and are important in the determination of dough extensibility (Békés *et al.* 2006). The LMW-GSs are encoded by the *Glu-A3*, *Glu-B3* and *Glu-D3* genes located on the short arms of the group 1 chromosomes (Sreeramulu *et al.* 1997). Additionally, three other loci have recently been identified, *Glu-2*, *Glu-4* and *Glu-5* on chromosomes 1B, 1D and 7D respectively (McIntosh *et al.* 2013). The LMW-GSs can be categorised as belonging to one of three biochemical groups based on their mobility under sodium-dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE): B, C and D types (Jackson *et al.* 1983), with the B type LMW-GS further divided among LMW-m, LMW-s and LMW-i according to the first amino acid residue, methionine, serine or isoleucine respectively (Muccilli *et al.* 2010). As with the HMW-GS, there is also allelic variation within the three LMW-GS encoding loci, with 17, 26 and 11 alleles reported for the *Glu-2*, *Glu-4* and *Glu-5* loci respectively (McIntosh *et al.* 2013).

1.5.2.2. *Gliadins*

Gliadins make up the the remaining 40-50% of wheat grain storage protein, and are generally poorly understood when compared to glutenins (Rasheed *et al.* 2014). Although they have a lesser impact than glutenins, gliadins also effect the processing and nutritional quality of the mature wheat grain (Khatkar *et al.* 1995; Rasheed *et al.* 2014). Whilst glutenins provide strength and elasticity to the dough, gliadins provide the dough with viscosity (Cornec *et al.* 1994; Khatkar *et al.* 1995).

The gliadins are a diverse group of monomeric proteins which can be divided into three distinct

groups: alpha- and beta-, gamma, and omega-gliadins ("The classification and nomenclature of wheat gluten proteins: A reassessment"). As with the glutenins, cysteine residues confer the ability for gliadin monomers to form intra-chain disulfide bonds (Shewry *et al.* 1997). However, unlike glutenins, they are unable to form inter-chain disulfide bonds, preventing the formation of large, heterogeneous high molecular weight molecules (Veraverbeke *et al.* 2002). The alpha-, beta-, and gamma-gliadins are both related to the LMW-GSs, contain six and eight cysteine residues respectively (Veraverbeke *et al.* 2002), and are classified as sulfur-rich prolamins ("The classification and nomenclature of wheat gluten proteins: A reassessment"). Omega-gliadins, however, are known as sulfur-poor prolamins due to their lack of cysteine residues and low methionine content ("The classification and nomenclature of wheat gluten proteins: A reassessment").

As with the glutenins, there is significant genetic variation of gliadin proteins present in the modern wheat population. All of the alpha- and beta- and some gamma-gliadins are encoded by the *Gli-1* loci on the short arms of the group 1 chromosomes, with 23, 24 and 15 alleles known for the *Gli-A1*, *Gli-B1* and *Gli-D1* loci respectively. Additionally, all omega- and the majority of gamma-gliadins are encoded at the *Gli-2* loci on the shorts arms of the group 6 chromosomes, with the *Gli-A2*, *Gli-B2* and *Gli-D2* loci having 36, 47 and 31 known alleles respectively.

1.5.2.3. *Non-gluten proteins*

Non-gluten proteins make up the remaining 15–20% of the protein in the wheat grain, and are mostly monomeric albumin and globulin proteins with either structural or metabolic roles (Goesaert *et al.* 2005). Such proteins include alpha-amylase, protease inhibitors, and enzymes with synthetic, metabolic, regulatory, or protective roles in the wheat plant (Singh *et al.* 2001b). In addition to the monomeric non-gluten proteins are the polymeric triticin globulin proteins, which play a minor storage role (Singh 1987). Although associations between some non-gluten proteins and bread-making performance have been made (MacRitchie 1987), the influence of non-gluten proteins on bread-making quality is poorly understood, and is generally considered somewhat insignificant in comparison to the far more abundant gluten protein (Veraverbeke *et al.* 2002)

1.5.3. Lipids

Often overlooked, lipids in the wheat grain fulfil an essential role in determining bread-making quality, largely due to their affinity for binding with starch and protein (MacRitchie 1987; Chung *et al.* 1978). The lipids present in wheat flour generally originate from cellular membranes and organelles. However, Hargin *et al.* (1980) identified triglyerides stored in spherosomes within the wheat endosperm.

Wheat grain lipids can be classified as either starch lipids, or free and bound non-starch lipids (Eliasson *et al.* 1990). The majority (66–75%) of total wheat flour lipids are free non-starch lipids, and are mostly triglycerides and other non-polar lipids. The bound non-starch lipids are

mostly associated with proteins, and consist of glyco- and phospholipids (Eliasson *et al.* 1990; Hoseney 1994). The starch lipids are primarily lysophospholipids, and are minor constituents of starch in the grain.

In relation to bread-making, starch lipids are so strongly bound to starch granules that they have little impact on dough quality (Goesaert *et al.* 2005). However, the polar lipids play an important role in increasing dough strength and handling characteristics, increasing gas retention, and improving loaf crumb structure (Graybosch *et al.* 1993; Gan *et al.* 1995).

1.6. Milling for bread-making

The purpose of milling is to produce flour from the wheat grain. For the production of white flour, milling is more specifically the process by which the starchy endosperm is removed from the aleurone and bran layers, and the embryo or germ (Posner 2009). In the UK, the majority of wheat is consumed as bread made from white flour, and it is the factors that affect the production of white bread-making flour that are of interest to this study.

Commercial wheat milling uses a series of rollers and sieves to break apart and grind wheat into different flours. The first stage of milling is the break, where whole grain are fed into rotating corrugated rollers, which tear the grain apart. This stage is repeated, and results in the production of some flour, but mostly various grain particles containing different combinations of the starchy endosperm, aleurone layer and bran, and embryo. These particles are then graded using a series of sieves based on the size and composition of the grain particle. Pure endosperm segments are directed into smooth rollers which reduce them into white flour, whilst other particles are sent for further sorting, or purification. During purification, grain particles are sorted based on their size, shape, and specific gravity using sieves and controlled airflow. Corrugated sizing rollers are then used to remove any remaining aleurone, bran, or embryo material adhered to the starchy endosperm, with the aim of complete separation of the endosperm. Extracted endosperm material is then directed to the smooth reducing rollers to produce white flour, whilst bran and germ by-products are processed separately. The result of this process is the production of multiple mill streams containing flour extracted from different parts of the grain and, due to gradients in the protein concentration and composition within the grain (discussed in section 1.7, with different protein compositions (Wang *et al.* 2007; He *et al.* 2013; Wan *et al.* 2014). The miller will then create a white flour that is a blend of different mill streams, and is suited to requirements of the production of the relevant food product. Likewise, the setting of the rollers can be adjusted to suit different grain, and to account for variations in grain size and hardness.

1.7. Protein gradients within the wheat grain

Distinct gradients have long been observed in the protein accumulated within mature wheat grain, and these gradients are now known to be both quantitative (Cobb 1905; Kent 1966; Tosi *et al.* 2011) and qualitative (Wang *et al.* 2007; He *et al.* 2013; Wan *et al.* 2014). These protein gradients have two notable effects on the end-use of the grain. The first of these effects is due

to the absolute gradient in protein, whereby gluten protein accumulates in the cells closest to the aleurone layer, with less protein present towards the central cells of the grain. During the production of white flour the aleurone and bran layers are removed from the starchy endosperm. However, the extraction rate (or milling yield) is not 100%, and the cells closest to the aleurone often remain adhered to the aleurone layer. This results in a disproportionate amount of protein being lost during the production of white flour. The second effect is caused by the differences in the accumulation of different gluten subunits within the endosperm (Wang *et al.* 2007; He *et al.* 2013; Wan *et al.* 2014). Since different mill streams are enriched with different parts of the grain, the result is the production of multiple flours each with a different protein composition, and with different baking characteristics (Yahata *et al.* 2006). However, to a certain extent this variation in mill stream quality can be accounted for by the miller when they mix the different flours to produce the final white flour product.

Whilst the presence of a protein distribution gradient has been known for over a hundred years (Cobb 1905), the mechanisms responsible for the formation of this gradient remain unknown. Recent work by Moore *et al.* (2016) used N^{15} labelled glutamine, the most abundant form of nitrogen transported into the developeing wheat grain (Fisher *et al.* 1986), to demonstrate that the precursors for protein production may be transported radially from the endosperm cavity, across the endosperm, to become concentrated in the subaleurone cells. The authors speculate that the subaleurone cells have a higher requirement for amino acids than the rest of the endosperm, and hence act as a strong sink, driving amino acid transport across the endosperm.

1.7.1. Quantification of protein gradients

Previous work on the protein gradients in the wheat grain have used techniques such as micro-dissection (Cobb 1905; Ugalde *et al.* 1990a; Ugalde *et al.* 1990b), pearl-milling (He *et al.* 2013) or milling on a laboratory experiment mill (Wang *et al.* 2007) and subsequent nitrogen and protein content analysis in mature grain, or sub-sampling of light- and immunofluorescence-microscopy images of developing grain (Tosi *et al.* 2011; Wan *et al.* 2014). These techniques have been relatively low-throughput, and so to facilitate the analysis of larger experiments with more combinations of treatment factors one of the aims of this project was to develop and implement a new high-throughput method to describe the protein gradients in the wheat grain.

1.8. **Future challenges for wheat production**

Over the next 80 years, it is predicted that the human population will increase by 50% (United Nations 2017). To support this population growth, it is estimated that food production will need to be increase by 50–100% as soon as 2050 (Southgate 2009; Parry *et al.* 2010). With wheat yields predicted to increase by only 38% over the same period (Ray *et al.* 2013), it seems unlikely that the required increase in production will be achieved through yield improvements alone. It should also be noted that these predictions are based on the extrapolation of the

yield increases achieved since 1989, and do not take into account the difficulties that will be faced in improving crop yields in the future. When combined with the pressures of policy and environmental change predicted for the future, it is likely that this deficit will be somewhat larger.

1.8.1. Reducing the environmental impact of agriculture

According to the Millennium Ecosystem Assessment, the greatest changes to terrestrial ecosystems have been due to the land-use changes associated with increased agricultural production (MEA 2005). Converting land for agricultural use disrupts both local ecosystems through habitat destruction, and also through the leaching of agrochemicals into watercourses, which can cause damage to both freshwater and marine aquatic life. To support the previous increased in food production, the amount of fertiliser applied has increased year-on-year, a trend that is likely to continue (Tilman 1999; Tilman *et al.* 2001). In addition to the on-going environmental damage caused by farming, the initial practice of clearing and tilling land for agricultural use is energy intensive, and results in the production of large amounts of carbon dioxide, further fuelling climate change (Tilman 1999). Tilman *et al.* (2001) predicted that a further 10^9 hectares of land will be converted for use in agriculture by 2050, an action that will cause "unprecedented ecosystem simplification, loss of ecosystem services, and species extinctions".

Fertilisation of agricultural land is associated with a range of environmentally damaging effects. The initial production of fertilisers is an energy-intensive process, and accounts for a third of the carbon emissions associated with crop production (Gellings 2009). The application of fertilisers (and other agrochemicals) is again an energy-intensive activity, producing further greenhouse gas emissions. Once applied the environmental damage continues. Of the nitrogen applied to agricultural land, as much as 2.5% is converted to nitrous oxide (Davidson 2009) (an ozone-depleting greenhouse gas with 300 times the global warming potential of carbon dioxide (Solomon *et al.* 2007)), which accounts for the majority of the relative greenhouse gas emissions from agriculture (Snyder *et al.* 2009). Whilst somewhat avoidable by proper management (Schröder *et al.* 2003), further environmental damage can be caused by the run-off and leaching of fertilisers into groundwater and watercourses, which may result in eutrophication and a decrease in biodiversity.

Whilst it is difficult to predict the future regulations that will govern global agriculture, it is likely that an increasing emphasis will be placed on reducing the environmental impact of food production. With regards to land-use, this could mean limiting the amount of land converted to agricultural land in an effort to protect natural ecosystems and biodiversity. As for the regulation of fertiliser usage, due to the energy-intensive nature of both the production and use of man-made fertilisers it seems likely that limitations will be imposed on the amount of fertiliser that can be applied to our soils in the future.

> The furnaces of the world are now burning about 2,000,000,000 tons of coal a year. When this is burned, uniting with oxygen, it adds about 7,000,000,000 tons of carbon dioxide to the atmosphere yearly. This tends to make the air a more effective blanket for the earth and to raise its temperature. The effect may be considerable in a few centuries.
>
> Francis Molena – *Popular Mechanics, March 1912*

1.8.2. Climate change

As a result of mankind's actions over the last two hundred years, earth's climate is changing. Until recently, the production and release of greenhouse and ozone-depleting gases has gone unchecked, resulting in a decrease in the levels of ozone in the stratosphere, and an increase in the greenhouse effect. As a result, global temperatures have increased, and will continue to do so (Alexander *et al.* 2006; Jenkins *et al.* 2009). In addition to a general trend towards higher temperatures, it is also predicted that the frequency of extreme weather events such as heatwaves and droughts will increase (Hennessy *et al.* 2008). In the UK, the most current predictions state that by 2080 mean maximum daily temperatures will rise by up to 5.4°C, the warmest day of summer will increase by up to 4.8°C, summer precipitation will reduce by up to 40%, and the frequency of droughts and heatwaves will increase (Jenkins *et al.* 2009). The effects of climate change will not be limited to summer, however, with average daily temperatures in winter also predicted to rise by up to 4.1°C. As a result, snowfall in the UK will be reduced by up to 95%. The effect of such dramatic changes to our climate will be to the detriment of agriculture and food security, as crops struggle to remain productive in increasingly hostile conditions (Wheeler *et al.* 2013).

Perhaps the most irrefutable, and widely known impact that human activity has had on the planet is the rapid increase in the levels of atmospheric CO_2 since the industrial revolution, a trend which will continue to an extent determined by the collective action of the international community (figure 1.5). Whilst the effects of greenhouse gas emissions on our climate have been known for decades (Sawyer 1972), the global community have been slow to act, with the first legally-binding global action to limit global warming made in the 2015 Paris Agreement. This agreement, signed by 195 nations, aims to limit global warming to a "safe" limit of 1.5°C, or to an absolute limit of 2°C. However, scepticism over the conclusions of research into climate change continues, with the United States of America (the second largest source of carbon dioxide emissions) intending to leave the Paris agreement in 2020 in favour of protecting the coal, oil, and gas industries (Zhang *et al.* 2017). Likewise, the Montreal Protocol to eliminate the production and use of ozone-depleting gases which has been enforced since 1989 is flaunted, and chlorofluorocarbons (CFCs) continue to be produced and released into the atmosphere (Montzka *et al.* 2018). Even assuming compliance with the emission mitigation policies, it is unlikely that the 2°C limit outlined in the Paris Agreement will be achieved (Brown *et al.* 2017; Raftery *et al.* 2017). Furthermore, even if fossil fuels were instantly eliminated altogether,

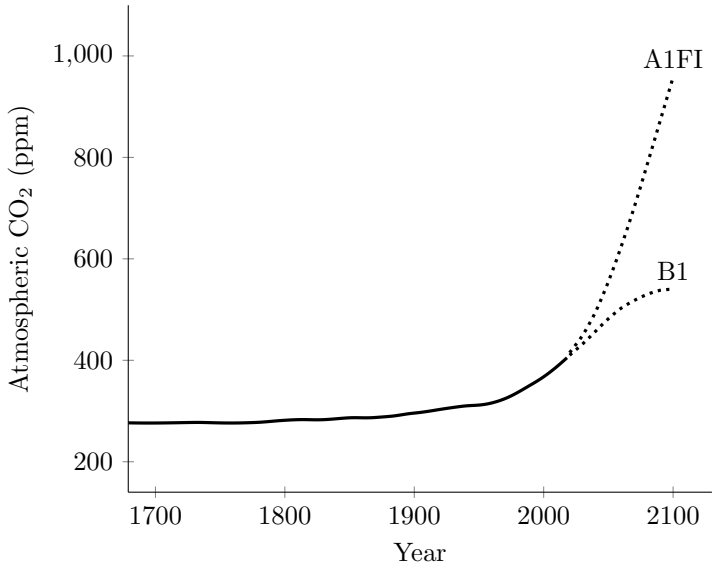


Figure 1.5: Historic levels of atmospheric CO_2, with predictions of future increases. Atmospheric CO_2 records from shortly prior to the industrial revolution to present day, as determined by ice core, firn, and direct atmospheric measurements of CO_2 by MacFarling Meure *et al.* (2006) (solid line). Future predictions (dotted line) of atmospheric CO_2 are presented based on the best- (B1) and worst-case (A1FI) scenarios outlined by the IPCC (2001). Both scenarios anticipate the human population to peak mid-century and to decline thereafter, with the A1FI scenario anticipating rapid economic growth supported by fossil-intensive energy supply, whilst the B1 scenario predicts shift towards a service- and information-based economy, with the introduction of environmentally friendly, resource-efficient technologies.

global warming will likely reach the 1.5°C "safe" limit target of the Paris Agreement through the residual action of past emissions (Mauritsen *et al.* 2017).

1.8.2.1. *The effect of climate change on wheat*

As with any crop, wheat is sensitive to changes in to its environment. Increased temperatures result in lower yields, with a drop of 6% for each °C that the temperature increases (Asseng *et al.* 2015). With the prediction that average summer temperatures in the UK will increase by up to 4.2°C (Jenkins *et al.* 2009), this indicates that we can expect a 25% decrease in wheat yield by 2080. However, this doesn't taking into account the disproportionate effect that heatwaves and droughts during grain-filling have on wheat production. Asseng *et al.* (2011) showed that sustained temperatures above 34°C during grain-filling can result in a 60% decrease in yield. In 1976, the UK experienced a heatwave in which the maximum daytime temperatures were above 30°C for 16 consecutive days. By combining the predictions that the summer maximum daily temperature in the UK to increase by up to 5.4°C (Jenkins *et al.* 2009) and that the frequency of heatwaves will increase (Hennessy *et al.* 2008), it is likely that the 60% decrease in yield observed by Asseng *et al.* (2011) will become a reality for British farmers unless adaptations are made to overcome such extremes in temperature.

The primary effect of high summertime temperatures on wheat is a reduction in the duration, but not the rate, of grain-filling (Altenbach *et al.* 2003; Gooding *et al.* 2003; Shah *et al.* 2003). The result of this is a decrease in the weight of the grain produced, which is mostly due to a reduction in the starch content of the grain. This effect was described further by Hurkman *et al.* (2011), who showed that high temperatures resulted in a decrease in the size of starch granules, which resulted in an increase in the relative proportion of protein in the grain. The authors also point out that these observations are known to result in variation in the quality of the flour produced from these grain. With regards to quality, high temperature is known to alter the protein composition of the mature grain, and is associated with a reduction in the relative amount of LMW-glutenins and omega-gliadins, and an increase in the proportion of HMW-glutenins and alpha-, beta-, and gamma-gliadins (Dupont *et al.* 2006b; Dupont *et al.* 2006a; Yang *et al.* 2011). However in contrast to the changes in grain yield, the changes in protein composition are generally positive, resulting in an increase in gluten content (Koga *et al.* 2015), protein and sulphur content, and Hagberg falling number Nasehzadeh *et al.* (2017). In the context of UK agriculture, this research suggests that there may be an increase in the amount of wheat meeting the quality requirements of bread-making. However, this will be accompanied by a decrease in both grain and milling yield (Dupont *et al.* 2006b). This decrease in milling yield is particularly relevant for the present study, which aims to describe how the differential concentration of protein in the outer endosperm is affected by factors such as increased temperature during grain-filling.

Whilst the majority of studies have focussed on the effect of increased summer temperature, the warming effect of climate change will apply year round. The predicted increase in winter temperatures is particularly relevant to UK wheat production, since the majority of wheat grown is winter wheat, and requires a period of vernalisation at low temperature over winter. The result of poor vernalisation is detrimental to the vegetative development of a wheat crop, resulting in a decrease in leaf emergence and tiller formation (Miglietta *et al.* 1995; Robertson *et al.* 1996).

The predicted increase in the frequency of heatwaves will be accompanied by more frequent droughts (Hennessy *et al.* 2008; Jenkins *et al.* 2009), with the two conditions often occurring simultaneously. The effect of drought on a wheat crop is a decrease in photosynthesis, leaf area, shoot mass, and grain yield (Shah *et al.* 2003), as well as inducing changes in grain protein content and composition (Altenbach *et al.* 2003; Yang *et al.* 2011). Furthermore, the effect of drought is often exacerbated when combined with an increase in temperature.

Among the numerous negative effects that climate change will have on wheat production in the future, the associated increase in atmospheric CO_2 (see figure 1.5) is likely to be beneficial. The comprehensive review by Amthor (2001) concludes that increasing the concentration of ambient CO_2 primarily results in an increase in grain yield, whilst also reducing the negative effects of drought. However, it also states that the positive effect of increase CO_2 may be

negated by moderate increased in ambient temperature. Crucially though, increased ambient CO_2 reduces grain protein content (either as a result of dilution caused by an increase in starch accumulation or a reduction in nitrogen uptake efficiency (Taub *et al.* 2008)) and decreases bread-making quality (Fernando *et al.* 2015).

1.9. **Aims and hypotheses**

1.9.1. Aims

- Develop and implement a high-throughput method for the quantification of protein gradients in the wheat endosperm.
- Identify the effect that elevated temperature during grain-filling, nitrogen supply, and genotype have on the protein distribution gradients in the wheat endosperm.
- Investigate how temperature, nitrogen supply, and genotype interact to alter the protein composition, and therefore bread-making quality of the mature wheat grain.
- Link the expression of gluten storage protein synthesis genes with any observed changes in protein accumulation during grain-filling.

1.9.2. Hypotheses

- Elevated temperature and increased nitrogen fertilisation increase the gradient in protein distribution in the wheat endosperm.
- Nitrogen input has a differential effect on the grain protein distribution gradients in different wheat varieties.
- Response to nitrogen input varies between genotypes and across years when grown in the field.
- Any factors affecting grain protein distribution also affect grain yield and yield components, protein composition, and the expression of protein synthesis genes.

Chapter 2: Materials and Methods

2.1. Introduction

To investigate the effects of temperature, nitrogen input, and genotype on grain morphology, protein distribution, protein composition, and related gene-expression, two practical experiments were conducted. To observe the combined effects of elevated temperature and reduced nitrogen supply, a single wheat variety, *Cadenza*, was grown in controlled-environment rooms with two post-anthesis temperature treatments, and two levels of applied nitrogen fertiliser. To investigate how nitrogen supply affects different commercial wheat genotypes in the field, four varieties grown under two levels of nitrogen were sampled over three years of the WGIN diversity field trial experiment.

Developing and mature grain was sampled from both the controlled-environment and field experiments for microscopy analysis of the protein distribution in the endosperm, protein composition according to SDS-PAGE and size-exclusion high-performance liquid chromatography (SE-HPLC), messenger ribonucleic acid (mRNA) expression related to protein production, and morphological characteristics of the grain. To investigate the differences in climate experienced between the three different years of field trials, meteorological measurements for temperature, accumulated thermal time, rainfall, and sunlight were recorded.

2.2. Controlled-environment post-anthesis temperature experiment

The controlled-environment post-anthesis temperature experiment was completed to investigate the combined effects of elevated temperature during grain-filling and limited nitrogen supply prior to anthesis on grain morphology and quality. Two different levels of nitrogen were supplied prior to anthesis, and plants were subjected to one of two temperature treatments after anthesis. The aim of this experiment was to simulate the effect that a prolonged heatwave during grain-filling might have on a wheat crop, both in terms of yield and quality, and how any negative effects may be exacerbated or alleviated by reduced nitrogen input.

2.2.1. Growth room conditions

The British spring wheat variety *Cadenza* was grown in controlled environment rooms at Rothamsted Research, Harpenden, UK in 2015. Two identical Weiss Gallenkamp controlled environment rooms were used for the experiment, with a floor area of 16m^2, growth area of 8m^2, and a height of 3m. Plants were sown five to a pot in nutrient-poor soil, and grown to anthesis in a single growth room under standard cereal growth conditions, supplied with either a high or a low nitrogen fertiliser regime, and with day/night temperatures of 20°C/15°C and humidity of 65%/75% respectively. The photoperiod was maintained at 16 hours throughout. After anthesis, which was determined as when three out of five plants per pot showed emerged anthers, half of the plants were moved to a second room with a higher daytime temperature of 28°C whilst maintaining the same night temperature of 15°C to prevent any unwanted

stress effects caused by elevated night-time temperatures (Prasad *et al.* 2008). Humidity and photoperiod remained unchanged at 65%/75% day/night and 16 hours respectively. The rooms were programmed to take two hours to change between the day and night temperatures to better reflect conditions seen in the field.

Lighting was provided by 400W HQI metal halide lamps with light intensity maintained at 500μmols m^{-2} sec^{-1} throughout. Additional tungsten lighting provided far-red illumination, with a negligible contribution to total light intensity. Light intensity was measured using a standard measuring point at the beginning and end of the experiment to monitor bulb performance. No bulbs required replacement during the course of the experiment, and there were no discrepancies found between the measured light intensity in each room at either the beginner or the end of the experiment.

2.2.2. Nutrient solutions

Nutrition was provided to the plants by liquid nutrient solution applied 12 times prior to anthesis. Two nutrient solutions were used, the high-nitrogen control, and a low-nitrogen solution containing $1/10^{th}$ of the nitrogen in the control. Plants were grown in the 'Rothamsted Nematode mix' provided by Petersfield Products (Leicester, UK). This potting mix consists of of 80% sterilised loam excavated from local construction sites (tested for contaminants, and enriched with a full-spectrum fertiliser), 15% sand and 5% 5mm lime-free grit. This potting mix was used to allow the effects of the different nutrient solutions to become apparent whilst providing a growth medium capable of facilitating normal plant growth without inducing undue stress. All pots were placed on saucers to prevent the nutrient solution from draining. During the course of the experiment, a total of 504 and 50.4mg of nitrogen was applied per pot under the high-and low nitrogen treatments respectively.

Nutrient solution was provided weekly until ear emergence, and biweekly thereafter. Nutrient solution was applied 12 times in total, with 500ml used per application. Plants were watered with reverse osmosis (RO) water as required; from six weeks onward, this was daily.

A 10× concentrated stock solution of each nutrient solution was made, and diluted immediately prior to application. One dilution was performed per experimental block of 32 pots to ensure any variation in the concentration of applied nutrient solution was contained within an experimental block, so that it could be accounted for in later statistical analysis.

The full composition of the nutrient solutions used are presented in table 2.1.

2.2.3. Experimental design

A complete randomised block design was used to grow 640 plants in 128 pots split equally between four experimental blocks, with a treatment structure of two levels of nitrogen fertiliser nested within two post-anthesis temperature regimes. Each experimental block contained 32 pots, of which 16 were subjected to either control or elevated post-anthesis daytime temperature. These 16 pots were again split between a high and a low nitrogen fertiliser

Table 2.1: Composition of the two nutrient solutions used in the controlled-environment experiment

Solution	Nutrient	Concentration
Macro-nutrients		
	KH_2PO_4	0.25 mM
	KOH	0.50 mM
	$MgSO_4.7H_2O$	0.75 mM
	$CaCl_2$	0.03 mM
	FeNaEDTA	0.10 mM
High N	$Ca(NO_3)_2.4H_2O$	4.00 mM
Low N	$Ca(NO_3)_2.4H_2O$	0.40 mM
Low N	$CaCl_2$	3.60 mM
Micro-nutrients		
	H_3BO_3	30.0 µM
	$MnSO_4.4H_2O$	10.0 µM
	$ZnCl_2.7H_2O$	1.0 µM
	$CuSO_4.5H_2O$	3.0 µM
	$Na_2MoO_4.2H_2O$	0.5 µM

regime, giving eight pots to be sampled at different timepoints during the experiment. Of these eight pots, two were sampled at harvest maturity (GS 92) to provide grain material for analysis and for yield calculations. Five of the pots were sampled at one of five timepoints during grain filling, and one pot (sampled at anthesis) was used in a separate study.

All plants were grown to anthesis in the same room under the low daytime temperature regime (see figure 2.1 for a schematic of the layout used). At anthesis, every other pot was removed in a chessboard pattern, and moved to the high-daytime temperature room, maintaining the original block structure. Figure 2.2 shows a schematic of the 128 pots used in the experiment as grown to anthesis in the first controlled environment room. Figures 2.3 and 2.4 show schematics for the 64 pots in the rooms used for the post-anthesis control and high temperature treatment respectively.

2.2.4. Sampling protocol

Grain was sampled at five timepoints during grain-filling, and again at maturity. The five timepoints used to sample during grain-filling (T1–5) were adjusted to account for the differences in accumulated thermal time between the control and high-temperature treatments. Material was also sampled at anthesis (T0) for use in another study. Material was collected from all four experimental blocks, but was only analysed from the first three blocks, since the additional level of replication proved unnecessary.

Sampling was conducted at set periods of thermal time after anthesis, as opposed to calendar days, as it allows for more robust comparisons to be made between plants grown under different temperature regimes. Without this approach, grain sampled from the high temperature treatment would be significantly more advanced developmentally than grain sampled from

Figure 2.1: Photo of Cadenza plants at 5 weeks in the controlled-environment room. A second identical room was used for the post-anthesis high temperature treatment.

the control temperature treatment.

Accumulated thermal time was calculated in growing degree hours (GDH) using hourly temperature figures from each of the two controlled-environment rooms, and a base temperature of 4.1°C. Values for the accumulated thermal time of each of five sampling timepoints for the control temperature treatment (10, 13, 21, 28, and 35 DPA) were used to calculate the sampling timepoints used in the high temperature treatment, with the result rounded to the nearest calendar day. Table 2.2 shows the sampling timepoints used in both DPA and thermal time (GDH).

Chlorophyll content of the flag leaf was monitored using a SPAD-502Plus meter (Konica Minolta, Tokyo, Japan) during the controlled-environment experiment to quantify the effect of the two nitrogen treatments. Measurements were taken 5 days after anthesis for all plants

Table 2.2: Sampling timepoints in DPA and GDH for each of the post-anthesis temperature treatments used in the controlled-environment experiment.

Timepoint (T)	Control (20°C)	High temperature (28°C)	Accumulated thermal time
0	0 DPA	0 DPA	0 GDH
1	10 DPA	7 DPA	3,440 GDH
2	14 DPA	10 DPA	4,816 GDH
3	21 DPA	15 DPA	7,224 GDH
4	28 DPA	20 DPA	9,632 GDH
5	35 DPA	25 DPA	12,040 GDH

Block I

1	Con N1 T4	*2*	High N1 T5	*3*	Con N1 T2	*4*	High N1 T3
5	High N2 T3	*6*	Con N2 M2	*7*	High N2 M2	*8*	Con N2 T1
9	Con N2 T2	*10*	High N1 T2	*11*	Con N1 T0	*12*	High N1 M2
13	High N2 T4	*14*	Con N2 M1	*15*	High N1 T0	*16*	Con N1 M2
17	Con N2 T4	*18*	High N2 T0	*19*	Con N1 T3	*20*	High N2 T2
21	High N2 T5	*22*	Con N2 T5	*23*	High N1 T4	*24*	Con N1 T1
25	Con N1 M1	*26*	High N2 M1	*27*	Con N2 T0	*28*	High N2 T1
29	High N1 T1	*30*	Con N1 T5	*31*	High N1 M1	*32*	Con N2 T3

Block II

33	Con N1 M1	*34*	High N2 T3	*35*	Con N2 M1	*36*	High N1 T3
37	High N1 T0	*38*	Con N1 T0	*39*	High N2 T4	*40*	Con N2 T5
41	Con N2 T2	*42*	High N2 T1	*43*	Con N1 T1	*44*	High N1 M1
45	High N1 T1	*46*	Con N1 T5	*47*	High N1 M2	*48*	Con N2 M2
49	Con N1 M2	*50*	High N2 T2	*51*	Con N2 T0	*52*	High N2 T5
53	High N2 T0	*54*	Low N2 T3	*55*	High N1 T5	*56*	Low N1 T3
57	Con N1 T4	*58*	High N2 M1	*59*	Con N2 T1	*60*	High N1 T4
61	High N2 M2	*62*	Con N2 T4	*63*	High N1 T2	*64*	Con N1 T2

Central walkway

Block III

65	Con N1 M2	*66*	High N2 T2	*67*	Con N1 T0	*68*	High N2 T1
69	High N1 T4	*70*	Con N1 T4	*71*	High N1 T0	*72*	Con N2 M1
73	Con N2 T1	*74*	High N1 T1	*75*	Con N2 T4	*76*	High N2 T4
77	High N1 T3	*78*	Con N2 M2	*79*	High N2 M1	*80*	Con N1 T3
81	Con N2 T2	*82*	High N2 T5	*83*	Con N2 T0	*84*	High N2 T0
85	High N2 T3	*86*	Con N1 M1	*87*	High N2 M2	*88*	Con N1 T2
89	Con N1 T1	*90*	High N1 M1	*91*	Con N2 T5	*92*	High N1 M2
93	High N1 T2	*94*	Con N1 T5	*95*	High N1 T5	*96*	Con N2 T3

Block IV

97	Con N2 T5	*98*	High N2 M2	*99*	Con N2 M1	*100*	High N1 T3
101	High N2 T0	*102*	Con N2 M2	*103*	High N2 T2	*104*	Con N1 M1
105	Con N1 M2	*106*	High N1 M2	*107*	Con N2 T3	*108*	High N1 T1
109	High N1 T5	*110*	Con N2 T4	*111*	High N1 M1	*112*	Con N2 T0
113	Con N1 T1	*114*	High N2 T1	*115*	Con N1 T0	*116*	High N1 T2
117	High N1 T0	*118*	Low N1 T4	*119*	High N2 M1	*120*	Low N1 T2
121	Con N1 T3	*122*	High N2 T5	*123*	Con N1 T5	*124*	High N1 T4
125	High N2 T3	*126*	Con N2 T1	*127*	High N2 T4	*128*	Con N2 T2

Door to room

Figure 2.2: Pre-anthesis experimental design used in the controlled-environment experiment. Con/High represent control of high temperature treatments and N1/N2 represent low and high nitrogen treatments respectively. T0–T5 represent the six sampling timepoints during development (see table 2.2), and M1/M2 the pots sampled at maturity. This design was maintained post-anthesis, with pots moved to a second room for the high temperature treatment in a chessboard pattern. Figures 2.3 and 2.4 show the post-anthesis layouts for the control and high temperature treatments respectively.

Block I

1	N1 T4	*2*		*3*	N1 T2	*4*	
5		*6*	N2 M2	*7*		*8*	N2 T1
9	N2 T2	*10*		*11*	N1 T0	*12*	
13		*14*	N2 M1	*15*		*16*	N1 M2
17	N2 T4	*18*		*19*	N1 T3	*20*	
21		*22*	N2 T5	*23*		*24*	N1 T1
25	N1 M1	*26*		*2*	N2 T0	*28*	
29		*30*	N1 T5	*31*		*32*	N2 T3

Block II

33	N1 M1	*34*		*35*	N2 M1	*36*	
37		*38*	N1 T0	*39*		*40*	N2 T5
41	N2 T2	*42*		*43*	N1 T1	*44*	
45		*46*	N1 T5	*47*		*48*	N2 M2
49	N1 M2	*50*		*51*	N2 T0	*52*	
53		*54*	N2 T3	*55*		*56*	N1 T3
57	N1 T4	*58*		*59*	N2 T1	*60*	
61		*62*	N2 T4	*63*		*64*	N1 T2

Central walkway

Block III

65	N1 M2	*66*		*67*	N1 T0	*68*	
69		*70*	N1 T4	*71*		*72*	N2 M1
73	N2 T1	*74*		*75*	N2 T4	*76*	
77		*78*	N2 M2	*79*		*80*	N1 T3
81	N2 T2	*82*		*83*	N2 T0	*84*	
85		*86*	N1 M1	*87*		*88*	N1 T2
89	N1 T1	*90*		*91*	N2 T5	*92*	
93		*94*	N1 T5	*95*		*96*	N2 T3

Block IV

97	N2 T5	*98*		*99*	N2 M1	*100*	
101		*102*	N2 M2	*103*		*104*	N1 M1
105	N1 M2	*106*		*107*	N2 T3	*108*	
109		*110*	N2 T4	*111*		*112*	N2 T0
113	N1 T1	*114*		*115*	N1 T0	*116*	
117		*118*	N1 T4	*119*		*120*	N1 T2
121	N1 T3	*122*		*123*	N1 T5	*124*	
125		*126*	N2 T1	*127*		*128*	N2 T2

Door to room

Figure 2.3: Post-anthesis experimental design used in the controlled-environment experiment for the control temperature treatment (20°C). See legend from figure 2.2 for explanation of notation.

Block I

1		*2*	N1 T5	*3*		*4*	N1 T3
5	N2 T3	*6*		*7*	N2 M2	*8*	
9		*10*	N1 T2	*11*		*12*	N1 M2
13	N2 T4	*14*		*15*	N1 T0	*16*	
17		*18*	N2 T0	*19*		*20*	N2 T2
21	N2 T5	*22*		*23*	N1 T4	*24*	
25		*26*	N2 M1	*2*		*28*	N2 T1
29	N1 T1	*30*		*31*	N1 M1	*32*	

Block II

33		*34*	N2 T3	*35*		*36*	N1 T3
37	N1 T0	*38*		*39*	N2 T4	*40*	
41		*42*	N2 T1	*43*		*44*	N1 M1
45	N1 T1	*46*		*47*	N1 M2	*48*	
49		*50*	N2 T2	*51*		*52*	N2 T5
53	N2 T0	*54*		*55*	N1 T5	*56*	
57		*58*	N2 M1	*59*		*60*	N1 T4
61	N2 M2	*62*		*63*	N1 T2	*64*	

Central walkway

Block III

65		*66*	N2 T2	*67*		*68*	N2 T1
69	N1 T4	*70*		*71*	N1 T0	*72*	
73		*74*	N1 T1	*75*		*76*	N2 T4
77	N1 T3	*78*		*79*	N2 M1	*80*	
81		*82*	N2 T5	*83*		*84*	N2 T0
85	N2 T3	*86*		*87*	N2 M2	*88*	
89		*90*	N1 M1	*91*		*92*	N1 M2
93	N1 T2	*94*		*95*	N1 T5	*96*	

Block IV

97		*98*	N2 M2	*99*		*100*	N1 T3
101	N2 T0	*102*		*103*	N2 T2	*104*	
105		*106*	N1 M2	*107*		*108*	N1 T1
109	N1 T5	*110*		*111*	N1 M1	*112*	
113		*114*	N2 T1	*115*		*116*	N1 T2
117	N1 T0	*118*		*119*	N2 M1	*120*	
121		*122*	N2 T5	*123*		*124*	N1 T4
125	N2 T3	*126*		*127*	N2 T4	*128*	

Door to room

Figure 2.4: Post-anthesis experimental design used in the controlled-environment experiment for the high temperature treatment (28°C). See legend from figure 2.2 for explanation of notation.

in the block, after all nutrient solution had been applied. Fifteen measurements were taken from each pot, with three replicated measurements from the flag leaf of each plant in a pot. Measurements were taken from pots for all sampling timepoints in the first experimental block only, resulting in 8 sets of measurements for each combination of temperature and nitrogen treatment.

For microscopy analysis of protein distribution gradients and protein body size-distribution, three grains were taken from the central third of the first ear to reach anthesis on one plant randomly selected from each pot. Transverse sections were cut in fixative from the middle of the grain, and stored in fixative prior to further processing (described in section 2.4). Samples were collected for microscopy analysis at five timepoints during grain-filling, but were only analysed from timepoints T2 and T3, since samples from the earliest timepoint (T1) contained insufficient protein to produce useful data, and samples collected at later timepoints (T4 and T5) could not be sectioned to a consistent standard. With regards to the timings of the cell enlargement stage of grain-filling (as discussed in section 1.3.1.2) the two sampling timepoints used, T2 and T3, likely represent early and mid to late grain-filling respectively.

For all other analysis, grain was sampled at five timepoints during grain-filling from the first three ears to reach anthesis on each of the four plants not sampled for microscopy. Sixteen grain were taken from the central third of each ear and immediately frozen in liquid nitrogen. Samples from each pot were bulked and ground for 120 seconds in liquid nitrogen using a SPEX SamplePrep 6870 freezer mill such that each sample contained material from twelve ears taken from four plants. Samples collected during grain-filling were analysed for nitrogen content (see section 2.6), protein composition by SDS-PAGE (see section 2.7), and mRNA expression analysis (see section 2.9).

Mature plants were hand cut and threshed, and grain samples bulked grain from all plants within a pot. Mature grain samples were oven dried to 5% moisture at 80°C. Measurements for yield and thousand grain weight were taken from grain sampled at maturity, and from these measurements an estimated grain count per plant was calculated. Mature grain was also analysed for nitrogen content (from which nitrogen yield per grain was calculated), protein composition by SDS-PAGE and SE-HPLC (see section 2.8), and for grain morphology. Grain area, length, and width measurements were obtained using the MARVIN grain analyser (GTA Sensorik GmbH, Neubrandenburg, Germany) from sub-samples of 100 mature grain from each treatment combination.

2.3. **WGIN diversity field trial experiment**

The WGIN diversity field trial is an ongoing long-term field experiment which has run since 2004 at Rothamsted Research, Harpenden, UK, in which 20–30 commercial wheat genotypes are grown under different levels of applied nitrogen fertiliser. Grain was sampled from this field trial in 2015, 2016, and 2017, with four varieties selected, grown under two different levels of nitrogen application. The aim of this experiment was to identify how grain-filling is affected by

different nitrogen inputs, different climate year-to-year, and how different genotypes respond to these factors.

2.3.1. Field conditions

Different fields hosted the WGIN diversity field trial in each of the three years of sampling used in this study, all within 2km of each other on the Rothamsted Research farm. In 2015, the WGIN field trial experiment was held in the *Bones Close* field, in 2016 in the *Blackhorse* field, and in 2017 in the *Great Harpenden* field. The *Bones Close* and *Great Harpenden* fields were the closest together, at approximately 600m, whilst *Blackhorse* was a considerable distance away, at 2km from *Great Harpenden*, and 1.5km from *Bones Close*. In 2015, 25 varieties were sown, whilst in both 2016 and 2017 there were 30 wheat varieties. The predominant soil type in each field was typical Batcombe (Avery *et al.* 1995).

Four levels of nitrogen fertiliser were applied to the WGIN diversity field trial: 0, 100, 200, and 350kg-N/ha. Nitrogen was applied in multiple applications, with an initial application of 50kg-N/ha for all treatments, a second application of 50, 100, or 250kg-N/ha, and a final application of 50kg-N/ha for the 200 and 350kg-N/ha treatments only. The timing of these applications, along with the timing of sowing, is shown in table 2.3. Soil nitrogen in the top 90cm prior to fertiliser application was measured at 36.5kg-N/ha in 2015, 48.0kg-N/ha in 2016, and 25.6kg-N/ha in 2017.

Of the four nitrogen treatments, two were sampled as part of this study: 100kg-N/ha and 350kg-N/ha. These two were selected to give samples with both an abundance of nitrogen and a moderate deficiency of nitrogen, whilst not severely stressing the plants.

2.3.2. Experimental design

A complete randomised block design was used in all three years to grow 25–30 wheat varieties under four different levels of nitrogen fertiliser, split equally between three experimental blocks. In 2015, 25 wheat varieties were grown, whilst in 2016 and 2017, 30 varieties were grown. The nitrogen fertiliser treatments were preserved between years, whilst the randomised layout of varieties and nitrogen treatments within each experimental block was changed each year.

Table 2.3: Drilling and nitrogen application timings for the three years of the WGIN diversity field trial experiment. First (N1) and second (N2) nitrogen application made to all treatments, and third (N3) application made to 200 and 350kg-N/ha only.

Year of experiment	Drilling	N1	N2	N3
2015 harvest	03/10/2014	16/03/2015	01/04/2015	30/04/2015
2016 harvest	12/10/2015	21/03/2015	08/04/2016	26/04/2016
2017 harvest	04/10/2016	15/03/2017	05/04/2017	09/05/2017

2.3.3. Wheat varieties sampled

The four wheat varieties sampled from the WGIN diversity field trial were selected to cover a range of phenotypes and end-uses. The varieties sampled were *Cadenza*, a British spring bread-making wheat (NABIM group 2), which is extensively used in glasshouse/controlled-environment experiments at Rothamsted Research; *Istabraq*, a low-protein high-yielding NABIM group 4 feed; *Hereward*, a benchmark NABIM group 1 bread-making wheat; and *Soissons*, a French bread-making wheat (NABIM group 2), which is photoperiod-insensitive (Bentley *et al.* 2013), and so anthesis occurs earlier than in other varieties, generally resulting in milder temperatures experienced during grain-filling. This phenotype may be particularly relevant in the future, since climate change is predicted to increase temperatures in both summer and winter, which will result in wheat plants reaching physiological maturity earlier in the year. Therefore, in the future an early-flowering wheat variety could be particularly desirable, since starting grain-filling earlier in the season could result in a reduction in the exposure of the developing grain to the highest temperatures of the summer.

Unfortunately due to technical issues, the final harvest of Soissons in 2017 was contaminated, and so no mature grain was available for analysis from Soissons in 2017.

2.3.4. Sampling protocol

Grain was sampled at five timepoints during grain-filling, and again at maturity. The five timepoints used to sample during grain filling were 10 DPA, 14 DPA, 21 DPA, 28 DPA, 35 DPA. The same timepoints were used throughout, with no correction made for differences in the thermal time experienced by different genotypes and between different years.

Sampling made during grain-filling was from 15 plants randomly selected in the middle of a 3m × 2.5m destructive sampling plot, with three plants sampled at each of the five sampling timepoints during grain-filling; sampling of mature grain was made from the adjoining 3m × 9m main plot, which was separated from the destructive sampling plot by 3m wide vehicle tracks.

For microscopy analysis of protein distribution gradients and protein body size-distribution, one grain was taken from the first tiller of three randomly selected plants at each sampling timepoint. Transverse sections were cut in fixative from the middle of the grain, and stored in fixative prior to further processing. Samples for microscopy analysis were analysed from a single sampling timepoint, T4 (28DPA). This timepoint was chosen as grain was at an approximately equivalent developmental stage to grain sampled at timepoint 3 in the controlled-environment experiment, and also represents the latest point (28DPA) point at which grain could be reliably sectioned and imaged. Additionally, it was only possible to analyse microscopy samples collected in 2015 and 2017, since samples collected in 2016 were destroyed due to malfunctioning equipment during sample preparation.

For all other analysis, grain was sampled at five timepoints during grain-filling from the same plants used for microscopy analysis, with 20 grain being taken from the centre of each of

the three ears sampled. Samples were bulked between the three ears sampled, immediately frozen in liquid nitrogen, and later ground for 120 seconds in liquid nitrogen using a SPEX SamplePrep 6870 freezer mill. All samples collected during grain-filling were analysed for nitrogen content and all Cadenza samples from the 2016 and 2017 experiments were analysed for RNA expression. Cadenza was chosen as the sole variety on which to complete RNA expression analysis due to its use in the controlled-environment experiment, and the years of 2016 and 2017 were chosen since the weather during grain-filling was markedly different between these two years, with a heatwave occurring in 2017 (see chapter 3).

The main plots were machine-harvested at maturity, and the grain samples bench dried to approximately 15% moisture. Measurements were taken for yield, thousand grain weight, and ears per plant, from which an estimated grain count per ear was calculated. Mature grain was analysed for nitrogen content (from which nitrogen yield per grain was calculated) and protein composition by SDS-PAGE and SE-HPLC.

2.3.5. Meteorological data

Meteorological data was collected from the on-site weather station at Rothamsted Research, Harpenden, UK (51.82°N, 0.37°W, 128m altitude) to monitor the climate during grain-filling over the three years of field trials (2015, 2016, and 2017). Measurements were taken for hours of sun, rainfall, and mean temperature, from which accumulated thermal time was calculated using a base temperature of 4.1°C.

The proximity of the weather station to the field trial site was 1.1km in 2015, 2.6km in 2016, and 800m in 2017.

2.4. **Microscopy analysis**

Image analysis of light-microscopy sections was used to quantify both the distribution of protein concentration and the protein body size-distribution in the endosperm of developing grain samples. Developing grain samples collected from the controlled-environment experiment and WGIN diversity field trial were fixed, embedded, sectioned, stained, imaged, and analysed. Analysis of microscopy images was completed using a novel image analysis software technique developed with Adam Michalski (Wrocław University of Environmental and Life Sciences, Wrocław, Poland) (Savill *et al.* 2018).

2.4.1. Fixation, dehydration, and embedding

Developing grain samples were fixed, dehydrated, and embedded in LR white resin to facilitate sectioning, staining, and imaging (protocol presented in appendix A.1).

Developing grain were removed from the ear by hand, and placed immediately in 4% (w/v) paraformaldehyde + 2.5% (v/v) gluteraldehyde fixative in 0.1M Sorensen's phosphate buffer (pH 7.4). Thin transverse sections of approximately 1mm were cut, and placed into glass vials of fixative. A light vacuum was applied and released three times for approximately two minutes each time, to aid the infiltration of fixative and removal of air from the grain sample.

Vials were placed in a sample rotator at room temperature for 3–5 hours, and then washed three times in 0.1M Sorensen's buffer (pH 7.4) for 30min in each wash. Samples were then stored at 4°C prior to further sectioning.

Samples were dehydrated in a graded ethanol series from 10–100% ethanol in steps of 10%. Samples were placed in a rotator for one hour for each ethanol concentration, and ethanol concentrations of 50–90% were repeated once. Samples were stored overnight at 4°C in 70% ethanol if required. Finally, samples were placed in 100% absolute ethanol for one hour, and repeated twice for a total of three hours in 100% ethanol.

Once dehydrated, samples were immediately infiltrated with medium grade LR white resin in a series of 4:1, 3:2, 2:3, and 1:4 ethanol/resin mixes. Samples were placed in each ethanol/resin mix on a rotator at room temperature for a minimum of six hours, and stored overnight at 4°C. After the graded series, samples were placed in 100% LR white resin for one hour, which was repeated twice to remove all ethanol. Samples were then stored in 100% LR white resin for at least five days, with the samples at room temperature on a rotator during the day, and stored at 4°C overnight. Resin was changed twice per day.

After infiltration with LR white resin, samples were polymerised so that the samples were contained in a solid resin block. Grain samples were carefully removed from the glass vials with tweezers and placed in pre-filled polyethylene capsules, ensuring the samples were sitting flush against the bottom of the capsule. Capsules were placed in a nitrogen-filled oven at 55°C for 16-24 hours until cured. Embedded sections were then left to cool prior to trimming and sectioning.

2.4.2. Sample sectioning, staining, and imaging

To image the wheat grain endosperm, embedded wheat grain samples were trimmed, sectioned, stained, mounted, and imaged with a light microscopy.

Resin blocks containing samples were trimmed down using a fresh double edged razor blade to create a trapezoidal cutting face. This reduces pressure on the microtome knife, and helps to produce a clean cut without striations. A diamond Diatome histology knife (Diatome Ltd, Nidau, Switzerland) was used to cut 1 µm sections on a Reichert Ultracut ultramicrotome (Ametek Reichert Technologies, Depew, NY, USA). Samples were floated on a bath of deionised water, removed using a fine paintbrush and placed onto deionised water drops on an eight-well glass microscopy slide (Hendley-Essex Ltd, Loughton, Essex, UK). Slides were dried on a hot plate at 80°C before staining.

Sections were stained for 30 seconds at room temperature with 1% (w/v) Naphthol Blue Black in 7% (w/v) acetic acid. Slides were then rinsed with deionised water and air dried prior to mounting.

Samples were mounted under no. 1.5 (0.17mm) glass cover slips with DPX Mountant to minimise distortion of images due to both starch bodies refracting light and from defects in

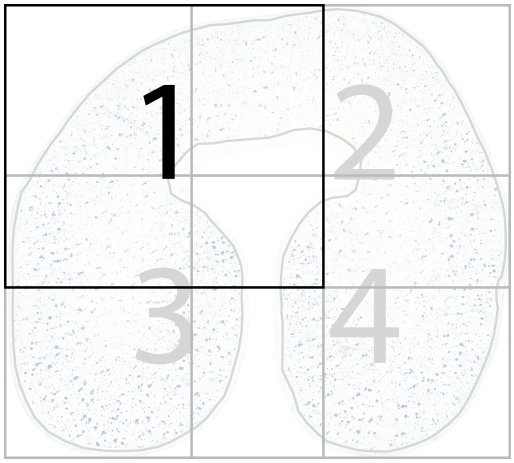


Figure 2.5: Diagram of microscopy image capture technique of four overlapping images per grain to ensure entire image was in focus. Individual images (labelled 1–4) were taken, and later combined into a single composite image prior to analysis.

the section.

Samples were imaged at 20× magnification using a Zeiss Axiophot light microscope (Carl Zeiss AG, Oberkochen, Germany). The MetaMorph (Molecular Devices, Sunnyvale, CA, USA) software package was used to automatically scan and image each section. Multiple images were mosaicked using the MetaMorph software to produce high resolution images. Four overlapping scans were taken per image, each covering different areas of the section. The microscope was re-focused prior to taking each scan. This approach was required to capture in-focus images, since any variation in slide angle or mountant thickness would cause the image to go out of focus as the microscope scanned the full width of the section; out-of-focus areas proved particularly problematic in the development of the image analysis technique, as small protein bodies would be obscured, and therefore not detected. Each of the four scans covered 62.5% of the width and height of the final image, as illustrated in figure 2.5. These four images were then checked for focus, and combined into a composite image using the 'photomerge' feature in Adobe Photoshop (Adobe Systems Inc., San Jose, CA, USA) with care taken to minimise blurred sections in the final image. A Python application was written to calculate the input coordinates of each of the four images (see appendix B).

2.5. **Microscopy image analysis**

A novel image analysis software technique was developed in ArcMap to detect and describe the pattern of protein within the wheat grain endosperm. The analysis uses a high-resolution light-microscopy image and a manually-drawn outline as inputs, and produces data describing both the gradient in protein concentration from the outside of the grain to the inside, and the size-distribution of protein bodies across the whole endosperm (Savill *et al.* 2018).

A prototype analysis toolbox was created using the ArcMap, part of ArcGIS 10.4 (ESRI®, Redlands, CA, USA), 'ModelBuilder' tool. This initial toolbox was used for the analysis of

the controlled-environment experiment data. Prior to the analysis of the WGIN field trial experiment data, the toolbox was re-written in Python, available online at doi: 10.5281/zenodo.1066914 (and in appendix C). The Python toolbox was functionally identical to the initial toolbox, but offered performance improvements, and introduced the automatic calculation of zone widths (explained further in section 2.5.1), conversion from pixels to micrometers, and improved formatting of the output data.

Both the protein concentration gradient and protein body size-distribution image analysis techniques rely on the supervised maximum-likelihood image classification method to identify areas of protein within the input image. For image classification to accurately detect areas of protein within a microscopy image, sectioned were stained with the selected protein stain Naphthol Blue-Black, which stains protein bodies dark blue and leaves the remaining areas largely unstained. This contrast between areas of interest and background is essential for the efficiency of the technique, as it relies solely on pixel colour and intensity for the detection of protein from background. For the supervised maximum-likelihood image classification, training sample areas must be defined for both protein and non-protein areas of the image. These sample areas must cover a range of pixel hues and intensities to correctly differentiate between protein and non-protein areas. The image classification protocol then compares these training samples against the image on a pixel-by-pixel basis, and each pixel is marked as either protein or background. Pixels marked as protein are then extracted from the background for measurement.

Microscopy images of stained wheat grain sections were loaded into ArcMap, and an outline manually drawn around the endosperm, just within the cells of the aleurone layer. One outline was drawn per grain image, and saved for use in the analysis. Training samples were then taken from each image, with ten samples selected from areas of protein, being careful to account for all variation present in the colour and intensity of the stained protein. Multiple samples were then taken that represented non-protein areas of the grain, including starch bodies, cell walls, and nuclei. Sufficient samples of non-protein were taken to ensure that all hues and intensities of pixels representing non-protein were taken into account. These training samples were then saved into an image classification file for use in the analysis toolbox. Three image classification files were created for each image, and the analysis run three times for each image using these classification files. This was to account for any variation in the training samples collected between images, since slight differences in the training samples selected inevitably alter the results of the analysis. The analysis toolbox was then run using the microscopy image, outline, and classification file as input. An overview of the analysis process is shown in figure 2.6.

2.5.1. Protein concentration gradient analysis

The protein concentration gradient analysis describes the changes in protein concentration from outer to inner endosperm, and relies on a zoning of the endosperm with five zones of

equal width drawn concentrically inwards from the outline of the endosperm. The width of each zone was calculated on a per-grain basis to account for differences in grain size. In the initial analysis toolbox used in the analysis of the controlled-environment experiment, a measurement was taken for the width of the two lobes of each imaged grain which was then halved, and divided to create five zones of equal width. In the python toolbox used in the analysis of the WGIN diversity field trial experiment, this process was fully automated. The resultant zones represent layers of the endosperm ranging from directly below the aleurone layer (zone 1), to the centre of the lobe (zone 5). Five zones were chosen as a result of preliminary empirical modelling on a sub-sample of data, which indicated that five zones were optimal for describing the distribution of detected protein.

The calculated zones were overlaid onto the extracted protein data, and area measurements for both the protein within each zone, and the total area of that zone recorded. From this data, a value for percentage protein by area was calculated for each zone, providing a profile of protein concentration from outer to inner endosperm tissues.

2.5.2. Protein body size-distribution analysis

The protein body size-distribution analysis was developed to investigate how the size of individual protein bodies varies in relation to their distance from the aleurone layer. It doesn't rely on the zones used in the protein concentration gradient analysis (section 2.5.1), but rather uses an exact measurement of distance from the aleurone layer. The analysis technique identifies protein bodies as groups of pixels representing protein (as determined by image classification), and measures the area of the protein body. It then measures the euclidean distance from the centre of the protein body to the outline drawn around the grain at the aleurone layer. Hence, it treats every single protein body as an individual datapoint, assigning an area and distance measurement to each one. This analysis is run concurrently with the protein concentration gradient analysis, but outputs its results separately.

Since individual pixels are often incorrectly identified as areas of protein, a limit was imposed which restricted the minimum size of each protein body to three pixels. This was deemed to be sufficient to still account for the smallest protein bodies, whilst excluding the majority of incorrectly assigned pixels. Such a consideration was not required for the protein concentration gradient analysis since single pixels identified as protein account for very little area, but were numerous enough to significantly affect the reliability of the protein body size-distribution analysis results.

2.5.3. Image analysis analysis data processing

Data from the microscopy image analysis methods had to be processed to improve the accuracy of the results, and to format the data for statistical analysis.

To improve the accuracy of the results from the protein concentration gradient analysis, a conversion was applied on the collected data using nitrogen content measurements taken from grain sampled from the same pot/plot at the same timepoint as the imaged grain. The

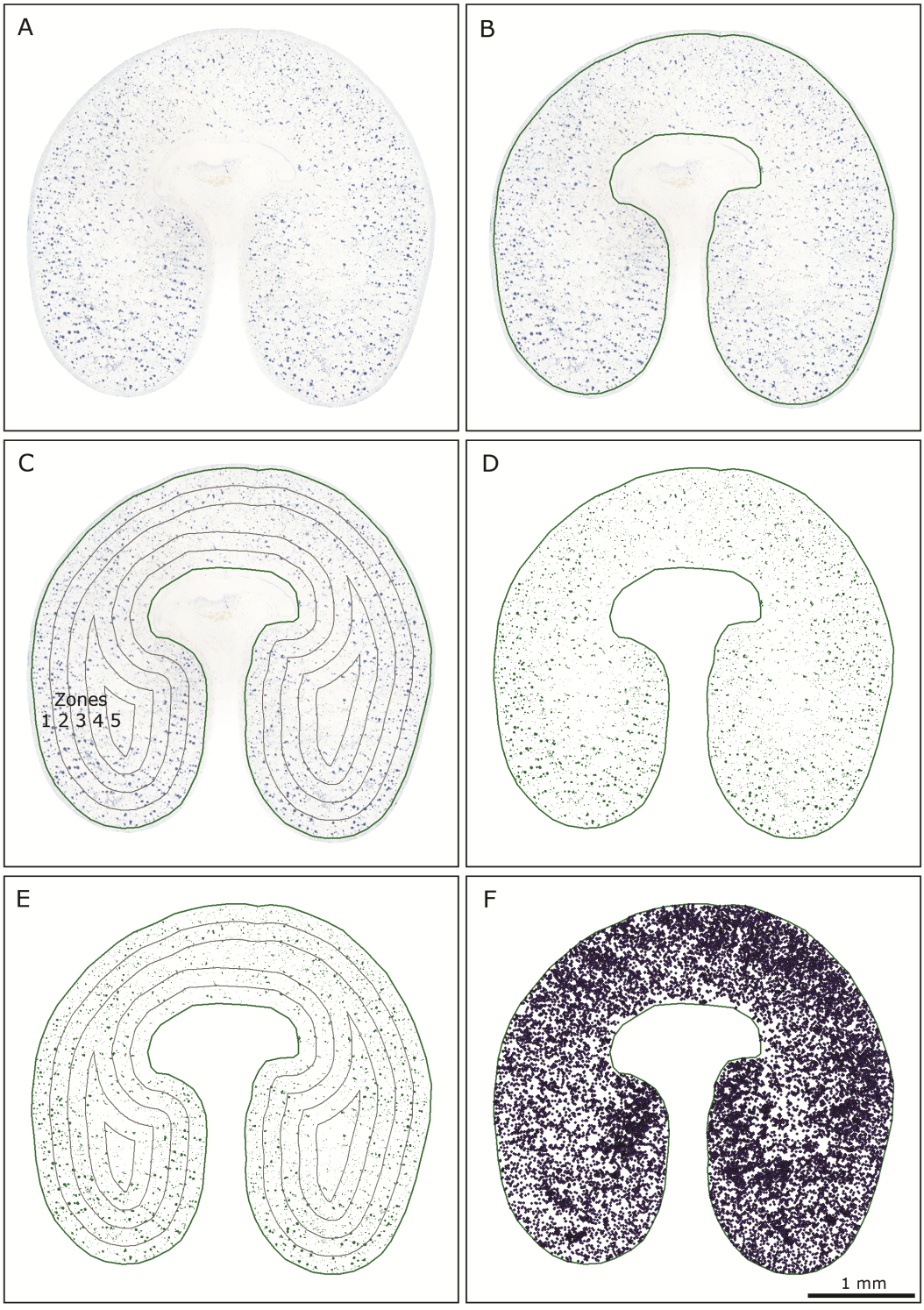


Figure 2.6: Grain protein composition image analysis workflow: top image shows the original wheat grain section, stained for protein with 1% Naphthol Blue Black in 7% acetic acid; middle left shows image with outline manually drawn around endosperm, just within the aleurone layer; middle right shows extracted protein data from image classification protocol; bottom left shows the zoning applied to describe the gradient in protein content; bottom right shows the final analysed image of detected protein overlaid with zones to calculate protein concentration by area for each zone.

nitrogen content data was multiplied by a factor of 5.7 (Sosulski *et al.* 1990), and used to calculate a conversion factor to correlate the grain protein content as detected by image analysis with the actual grain protein content. A unique conversion factor was calculated for each treatment combination, that effectively normalised the amount of protein detected by image analysis. This was deemed necessary since initial analysis showed some disparities between protein detected by image analysis, and actual protein measurements made on the grain. It is hypothesised that these differences are caused by differences in protein density, which are not taken into account by our by-area measurements of protein concentration. The formula for the calculation of the conversion factors applied was as follows:

$$\text{Conversion factor} = \frac{\text{Protein concentration from nitrogen content measurements}}{100\times(\frac{\text{Total protein area from microscopy image analysis}}{\text{Total grain area from microscopy image analysis}})}$$

This conversion was applied to all data collected from the protein concentration gradient analysis method, from both the controlled-environment and WGIN diversity field trial experiments. No conversion was applied to the protein body size-distribution data, since the protein body size itself is of interest rather than the amount of protein. However, the data from the size-distribution analysis was filtered to removed erroneous measurements: any protein bodies detected below an area of three pixels were removed, since during testing objects below this size threshold rarely represented correctly identified protein bodies.

Whilst the collection of protein concentration gradient data from the controlled-environment experiment was performed manually, the Python toolbox used for the analysis of the WGIN diversity field experiment automatically output data as comma separated values (CSV) files, with a single file produced for each run of the analysis. To quickly collate these individual results into a single dataset file that could be used for statistical analysis a Python script was written, and is presented in appendix G.

Similarly, the protein body size-distribution analysis method outputs single CSV files, each containing individual measurements of the area, and distance from the aleurone layer of every protein body detected within the image. Due to the number of data files produced, and the size of these files, it is impossible to manipulate the data using traditional spreadsheet software, and so a Python script was written to combine the CSV files, and also to assign each protein body to one of the five zones (as used in the protein concentration gradient analysis) which was used in the production of figures, but not for any data analysis. The script presented in appendix D was used to produce a list of the maximum grain width (from aleurone layer to the centre of the lobe of the grain), which was used to attribute size-distribution datapoints to the relevant endosperm zone. The output of this script was then used with scripts to join the multiple CSV files produced by the size-distribution analysis. The Python scripts used to collate the results from the controlled-environment and WGIN diversity field experiment are presented in appendices E and F respectively.

2.6. Nitrogen content analysis

Nitrogen content of wholemeal flour was determined by the Dumas method using a LECO CN628 Combustion Analyser (LECO corporation, St Joseph, MI, USA) on 0.5–1g of freeze-dried ground grain material. This method measures nitrogen content as a percentage of dry matter, from which protein content was calculated using a factor of 5.7 (Sosulski *et al.* 1990).

2.7. SDS-PAGE analysis

SDS-PAGE analysis was used to measure the relative abundance of the reduced gluten protein subunits within the wheat grain samples collected during this study. Gluten protein was selectively extracted, and then run on electrophoresis gels to separate the individual protein subunits. Images of the stained gels were then captured and analysed.

2.7.1. Protein extraction for SDS-PAGE

Prior to protein extraction all grain samples were homogenised for 120 seconds in liquid nitrogen using a SPEX SamplePrep 6870 freezer mill. Once milled, the whole-grain flour was kept frozen in liquid nitrogen, and sub-samples were taken and freeze-dried. Dried flour samples were kept in sealed plastic tubes at room temperature until use. The extraction protocol used was adapted from the method for extraction and separation of wheat gluten proteins described by Tatham *et al.* (2000), and is presented in appendix A.2.

Wheat gluten storage proteins were extracted from 10mg flour samples in 150µl propan-1-ol + 2% dithiothreitol (DTT), to act as a reducing agent to break down the inter-chain disulphide bonds of the gluten proteins, in a heated shaker at 50°C for 45 minutes. Samples were centrifuged at maximum speed for 15 minutes, and the supernatent removed and retained. The pellet was resuspended in 150µl propan-1-ol + 2% DTT, and placed in a heated shaker at 50°C for a further 45 minutes. Samples were centrifuged at maximum speed for 15 minutes, and the supernatents combined. The extracted protein was freeze-dried overnight. Dried protein was resuspended in 150µl of total loading buffer: 2% (w/v) sodium-dodecyl-sulphate (SDS), 200mM DTT, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue in 50mM Tris-HCL (pH 6.8). Protein samples were then heated to 90°C for three minutes, and centrifuged at maximum speed for 15 minutes. The supernatent containing extracted gluten protein was then run on an SDS-PAGE gel. Extracted protein was stored at -20°C, and heated at 90°C for three minutes prior to use.

2.7.2. SDS-PAGE

Wheat gluten protein samples were separated using the Bolt® Mini Gel Tank electrophoresis system. This system utilises pre-cast Bis-Tris gels to obtain consistent results both across and between runs. Stained gels were then imaged and analysed. The full technical protocol is available in appendix A.3.

Bolt® pre-cast 8% Bis-Tris gels with 17 wells were used throughout the experiment to separate protein samples for analysis. Gel cassettes were rinsed in deionised water, the wells washed

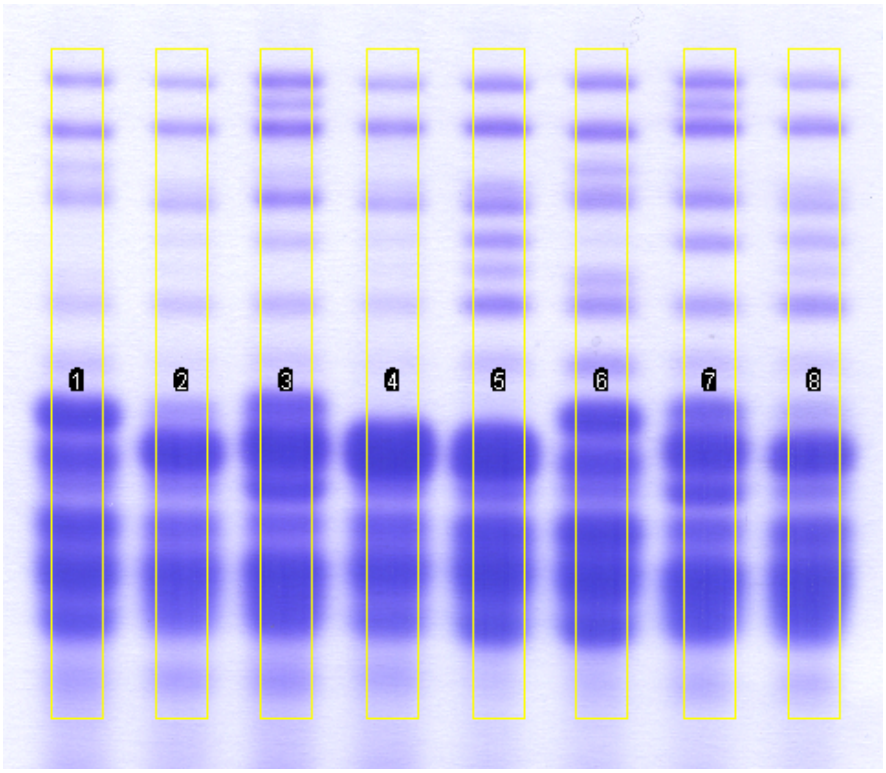


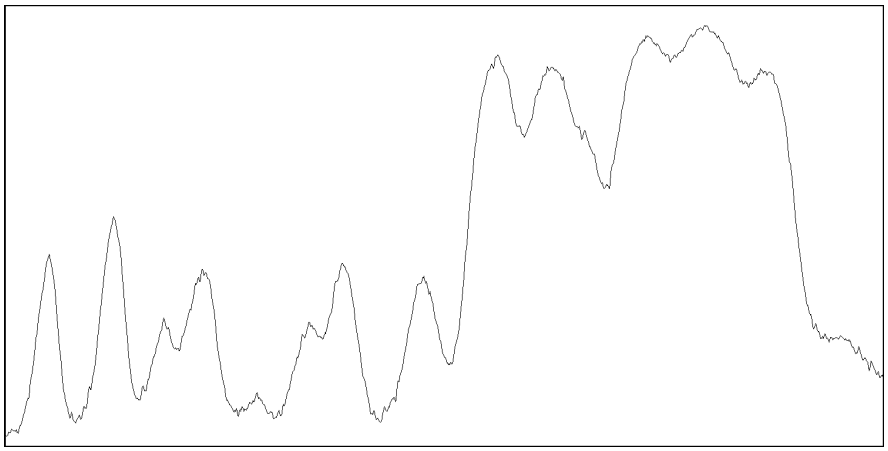
Figure 2.7: Image showing lane selection for protein gel analysis in FIJI. A rectangular lane is drawn over than protein bands of interest, and copied on to each lane. The intensity of each band within each lane is then measured.

with MES running buffer, and then fitted into the gel tanks. Six gels were run across three tanks for each of the two experiments, to ensure that separation was consistent for all samples in each experiment. The gel tanks were filled with MES buffer, and 1–10µl of protein sample loaded per well (adjusted through trial-and-error to produce the clearest separation). Gels were run at a constant current of 20mA for 30 minutes to settle the sample in the well, and then at 50mA for a further 330 minutes. Gels were then removed from their cassettes and placed in Coomassie stain overnight: 0.1% (w/v) Coomassie Brilliant Blue (R-250), in 40% (v/v) methanol with 10% (w/v) trichloroacetic acid (TCA). Stained gels were then destained in 10% (w/v) TCA, with a small strip of foam in the container to absorb the dye, until the background was clear. Gels were then imaged using a flatbed scanner prior to analysis.

2.7.3. SDS-PAGE image analysis and data processing

Images of the stained gels were captured at 1200dpi using an HP Scanjet G4010 flatbed scanner, and stored as 24-bit TIF files. Captured images were analysed using the gel analysis feature of the open-source software package FIJI (Schindelin *et al.* 2012). This analysis measures the intensity of bands running down each lane of the gel, and can be used to compare relative protein levels between treatments. Identical rectangles were drawn onto each of the lanes of the gel, as shown in figure 2.7. Once all lanes were selected the analysis was run and intensity levels for each lane generated. The output of the analysis of each lane on the gel is

(a)



(b)

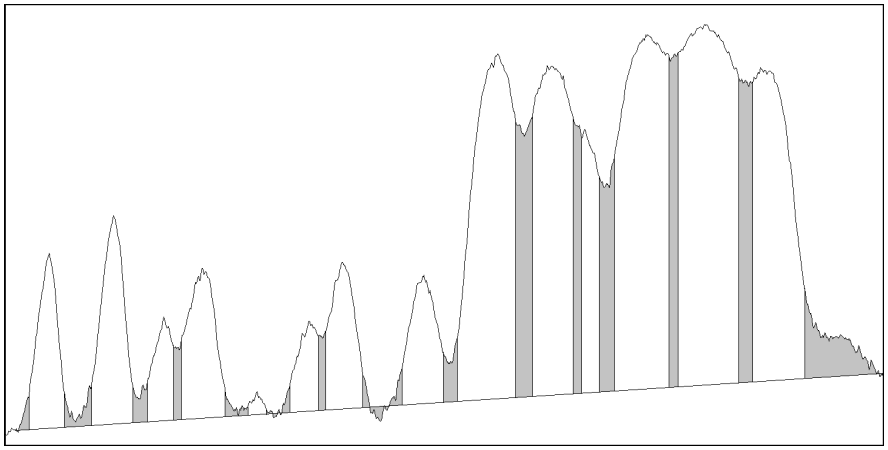


Figure 2.8: Images showing output of protein gel analysis in FIJI. Sub-figure a shows the raw pixel intensity output from one lane of the gel analysis in FIFI (lane 6 from figure 2.7). Sub-figure b shows the pixel intensity graph with a baseline ruled along the bottom of the peaks to remove background staining, and vertical ruling of the limits of each detected band in a lane. The area of these peaks is then measured, and this absolute value is compared against the total area under all peaks to give a relative measurement of each protein band.

a continuous graph of pixel intensities within the selected rectangle, as shown in figure 2.8a. The peaks of these graphs were ruled across at the base to remove background information. The area of the remaining ruled off peak represents the pixel intensity of the relative band on the protein gel. Figure 2.8b shows the output graph from figure 2.7 with each individual peak ruled off. The area under each peak was then measured using the wand tool in FIJI, and the pixel intensity values recorded.

Areas under peaks in the pixel intensity output graphs were recorded, and represent the relative abundance of a particular gluten protein subunit within the total gluten protein extract. Due to the fact that that the efficiencies of the protein extraction and subsequent staining on the electrophoresis gel, absolute measurements were not used in the analysis. Relative levels of each protein subunit with respect to the total amount of gluten protein detected were calculated, and used for statistical analysis.

2.8. **SE-HPLC analysis**

SE-HPLC analysis was used to quantify the abundance of different protein groups in the wheat grain, to give an approximation of baking quality in lieu of any direct measurements of flour performance. Whilst SDS-PAGE measures the relative amount of reduced proteins present in a flour sample, SE-HPLC analysis quantifies protein polymers. As such, SE-HPLC analysis provides a lower resolution approach to measuring grain composition, providing measurements of five protein fractions enriched with either HMW-glutenins (F1), LMW-glutenins (F2), omega-gliadins (F3), alpha-, beta-, and gamma-gliadins (F4) or albumin and globulins (F5).

2.8.1. Protein extraction for SE-HPLC

Protein extracts for SE-HPLC were made from whole-grain flour that had been homogenised and freeze-dried as described in section 2.7.1. Total protein extracts were prepared using sonication with SDS in a phosphate buffer using a protocol (presented in appendix A.4) adapted from the method described by Millar (2003).

Wholemeal flour sub-samples of 16.5mg were extracted in 1.5ml of SDS protein extraction buffer: 2% (w/v) SDS in 0.1M NaH_2PO_4 buffer, corrected to a pH of 6.9 with HCl. Samples were sonicated for 45 seconds with an ultrasonic disintegrator fitted with a 3mm exponential tip, and an amplitude of 6µm. Samples were then centrifuged at maximum speed for 30 minutes and the supernatant removed. The protein extract was filtered with a 0.45µm filter and transferred into 2ml glass vials prior to analysis by SE-HPLC.

2.8.2. SE-HPLC analysis

Grain protein extracts were analysed using a Shimadzu (Kyoto, Japan) SE-HPLC machine with a Phenomenex BioSep® 5µm SEC-s4000 column and a 50% (v/v) acetonitrile with 0.1% (v/v) trifluoroacetic acid (TFA) eluent. 20µl of each sample was analysed with a flow of 0.2ml per minute for 25 minutes. A detection frequency of 210nm was used to quantify the abundance of protein fractions in each sample.

2.8.3. SE-HPLC data collection and processing

Data collected from the SE-HPLC analysis of wheat grain protein extracts was in the form of a ultra-violet (UV) (at 210nm wavelength) absorption curve. The area underneath the peaks of this graph provide quantification of five protein fractions enriched with either HMW-glutenins, LMW-glutenins, omega-gliadins, alpha- and gamma-gliadins, or albumin and globulins. Figure 2.9 shows an example absorption curve adapted from Millar (2003).

Identification of the five peaks shown in figure 2.9 was semi-automated, and quantification of the area under each peak completed by the LCsolution software (Version 1.22, Shimadzu Corporation, Kyoto, Japan). Results were checked for correct identification of peaks prior to analysis.

As with the SDS-PAGE analysis, the proportion of each SE-HPLC protein fraction of the total detected protein was analysed. However, this was not due to potential variation in the efficiency of the protein extraction, but rather in the fact that the absolute protein measurements from the SE-HPLC analysis were heavily compounded with the overall increase in grain protein concentration. Therefore, analysis of the levels of different protein groups was made using relative measurement of each protein fraction as a percentage of the total protein detected. Additionally, the relative ratios of F1/F2 (the ratio between HMW-glutenins and LMW-glutenins), of (F3+F4)/F1 (the ratio between gliadins and HMW-glutenins), and of gluten protein to total protein were analysed.

Whilst the absolute measurements from the SE-HPLC analysis were not presented in isolation, absolute SE-HPLC values were used in the linear regression analysis with accumulated gene expression data (as described in section 2.10). This was deemed appropriate since this regression analysis was an exploratory approach to identify a link between the gene expression results (presented in chapter 7) with the amount of protein present in the grain. This analysis would not have been possible with relative protein measurements.

2.9. **RNA expression analysis**

Real-time polymerase chain reaction (PCR) (quantitative PCR (qPCR)) was used to analyse the expression of gluten protein synthesis genes, to identify how mRNA expression was affected by temperature and nitrogen supply in the controlled-environment experiment, and by different nitrogen fertiliser regimes in the WGIN diversity field trial experiment. For the WGIN diversity field trial experiment, expression analysis was only completed on *Cadenza* grain sampled during 2016 and 2017. This was to allow for comparison with the controlled-environment experiment, which also used *Cadenza*, and to compare between two years with a significantly different weather during grain-filling.

2.9.1. Primer selection

For qPCR, oligonucleotide primer pairs were designed to amplify all homeologous copies of the gluten protein synthesis genes for alpha- and beta-gliadins, gamma-gliadins, omega-gliadins, HMW-glutenins (A- and B-type), and LMW-glutenins. Primers were also designed for the 18S

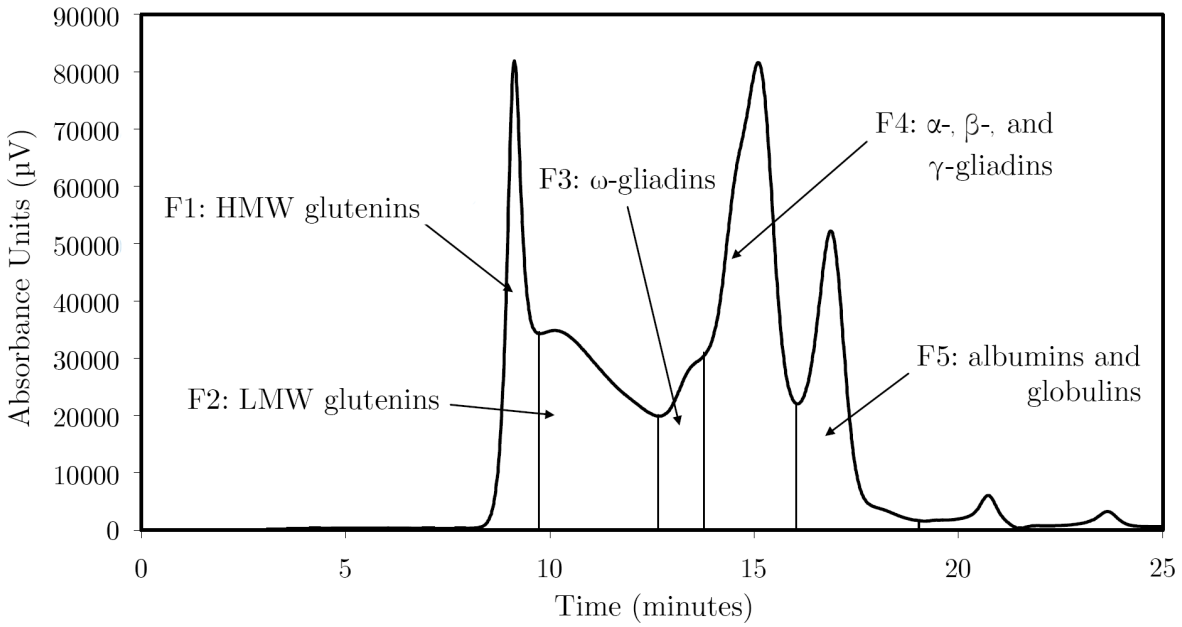


Figure 2.9: Graph of example SE-HPLC absorbance curve for wheat grain protein extract, with labelled peaks corresponding to the five protein groups detected. Adapted from Millar (2003).

and 28S mRNA reference genes (table 2.4). Primer pairs were designed by Dr Peter Buchner (Rothamsted Research, Harpenden, UK) with a length of 20-25 bases, GC content of 45-60%, T_m of 62–66°C (Nakano *et al.* 1999; Kibbe 2007), T_m differences of 5°C or less between pairs, and a maximum of two degenerate bases.

2.9.2. RNA extraction

Ribonucleic acid (RNA) was extracted from grain samples collected to analyse the expression of gluten protein subunit synthesis genes in grain subjected to different treatments. The extraction of RNA was performed on grain samples which were removed from the plant, immediately frozen with liquid nitrogen and milled in a freezer mill. Samples were not freeze-dried, and at no point were samples allowed to defrost prior to RNA extraction. RNA extraction was only completed on developing grain, and the full technical protocol is presented in appendix A.5 for reference.

Prior to RNA extraction, all relevant solutions were treated with 0.1% (v/v) diethylpyrocarbonate (DEPC) to destroy potential RNase activities. The solutions were thoroughly shaken, left to incubate at room temperature for several hours, and autoclaved prior to use.

RNA was extracted from approximately 0.5–1.0g of fresh, frozen milled grain in a hot (80°C) phenol/extraction buffer (8:12 ratio). The extraction buffer contained: 1% (w/v) SDS, 0.1M LiC, and 10mM EDTA in 0.1M Tris-HCl (pH 8.0). 1ml of hot phenol/extraction buffer was added to the frozen ground sample, vortexed until completely thawed and homogenised, and then vortexed for a further 30 seconds. 0.5ml of a 24:1 mix of chloroform/isoamyl alcohol (IAA)

Table 2.4: Sequences of primers used for real-time PCR analysis of gluten protein synthesis genes, and the 18S and 28S ribosomal RNA (rRNA) reference genes.

Primer	Sequence (5' to 3')	Amplicon size	T_m
18Ssqas2	CTG GTC GGC ATC GTT TAT GGT TG	125bp	64.6
18Srts1	AA CTG CGA AAG CAT TTG CCA AGG		65.2
28S-Ribo-rtFor1	GTT TAC GGC GAT GTT AGG AAG TC	95bp	62.9
28S-Ribo-rtRev1	AAC CGT TTC CAA GGT TGA CAG GC		64.6
Ta-abGliaFor3	ATG AAG ACC TTT CTC ATC CTT GCC	92bp	63.6
Ta-abGliaRev3	TGT GGC TGC AAT TGT GGC ACT G		64.2
TaGlia1For	CCT GCG GCC ACT ATT TCA GCT C	126bp	65.8
TaGlia2For	GCA GCA ACA GGT GGG TCA AGG T	132bp	65.8
TaGlia1/2Rev	GGG ACA TAC ACG TTG CAC ATG G		64.2
Ta-OmegaGliaFor1	TGT CCT CCT TGC CAT GGC GAT G	80bp	65.8
Ta-OmegaGliaRev1	TGA GGT GAT TGT ARC TCT TTG TTG C		62.5–64.1
TaLMW-Glut-abFor3	GAA GAC CTT CCT CRT CTT TGC C	96bp	62.1–64.2
TaLMW-Glut-abRev3	TGG TCT CTC CAA ACC AGG GAT G		64.2
TaHMW-ABrtFor1	GAG ATG GCT AAG CGC YTG GTC		63.2–65.8
TaHMW-ABrtRev1a	GCT CGC GCT CAC ACT GTA GTT G	105bp	65.8
TaHMW-ABrtRev1b	GCT CGT GCT CAC ATT GTA GTT GTC		65.2

was then added and vortexed for a further 30 seconds. The samples were then centrifuged at maximum speed for five minutes at 4°C. The aqueous phase was removed and transferred to a new 2ml micro-centrifuge tube. If the interphase was large, the initial extraction stage was repeated, and the aqueous phases combined. To the aqueous phase, 1ml of chloroform/IAA was added, vortexed for 30 seconds, and centrifuged at maximum speed for five minutes at 4°C. The aqueous phase was then transferred to a new 1.5ml micro-centrifuge tube. The volume was measured, and an equal volume +20µl of 4M LiCl was added. The samples were mixed well and incubated at 4°C overnight.

Incubated samples were centrifuged at maximum speed for 20 minutes at 4°C, and the supernatent discarded. The pellet was washed with 1ml of 70% ethanol, centrifuged at maximum speed for five minutes at 4°C, and the supernatent discarded. The pellet was then allowed to air dry. A DNAse treatment was prepared: 8µl RNAse-free DNAse, 15µl 10× DNAse buffer in 127 DEPC-treated H_2O per sample. 150µl of the DNAse treatment was added to each pellet, and dissolved on ice for 30–40 minutes. When completely dissolved, the samples were incubated at 37°C on a shaker for 30 minutes. DEPC-treated H_2O was added at 150µl per sample, and 300µl of chloroform/IAA was added and vortexed for 30 seconds. The sample was then centrifuged at maximum speed for five minutes at 4°C. The aqueous phase was transferred to a new 1.5ml micro-centrifuge tube and 300µl of phenol/chloroform/IAA added, and vortexed for 30 seconds. The sample was then centrifuged at maximum speed for five minutes at 4°C.

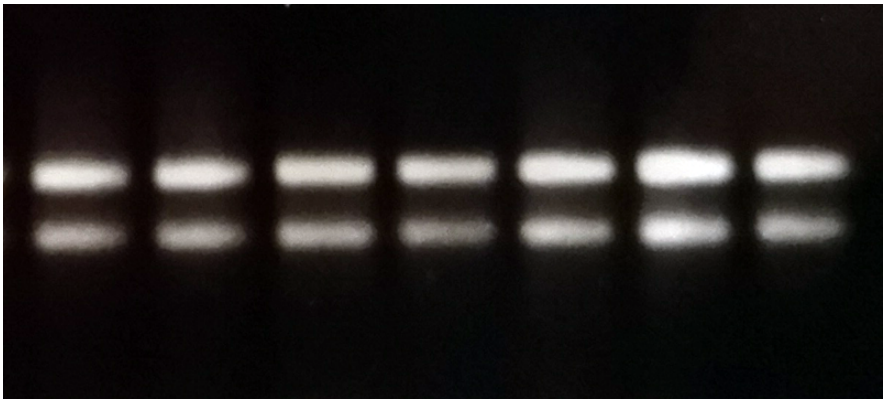


Figure 2.10: Image of example RNA quality check result. Two crisp bands indicate good quality RNA. The upper band represents the 28s rRNA (95bp), and the lower band is the 18s rRNA (125bp).

After a second extraction with chloroform/IAA, the aqueous phase was transferred into a 1.5ml micro tube, and the total RNA precipitated by adding 1/10 volume of 3M NaOAc (pH 5.2) and 2.5× volume of 100% ethanol. The extract was mixed well and incubated at -20°C overnight.

After centrifugation at maximum speed for 20 minutes at 4°C, the supernatent was discarded, and the pellet washed with 1ml of 70% ethanol by further centrifugation for five minutes at 4°C. The supernatent was discarded and the pellet allowed to air dry. The pellet was dissolved on ice in 30–150µl of DEPC-treated H_2O (depending on the size of the pellet). Once dissolved, the extracted RNA was heated to 37°C in a shaking heat block for five minutes and centrifuged at maximum speed for five minutes at 4°C. The supernatent was transferred to a fresh 0.5ml micro-centrifuge tube, and the concentration of RNA measured at 260nm using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Hemel Hempstead, UK). RNA was diluted with DEPC-treated H_2O to a concentration of 1000ng/µl or less.

The quality of extracted RNA (as measured with the NanoDrop 2000c spectrophotometer) was then tested by tris-acetate-EDTA (TAE)-agarose electrophoresis. A 1% (w/v) agarose in TAE gel was prepared with 1% (v/v) Sybr® Safe DNA Gel Stain for RNA detection. 1µg of RNA in 10µl water and RNA loading dye was loaded per well. The gel was run for 40 minutes at a constant 60V and observed under UV light. An example result of this quality check is shown in figure 2.10

2.9.3. cDNA synthesis

Complementary DNA was synthesised from the extracted total RNA, to be used in qPCR analysis (full protocol in appendix A.7). Using the RNA concentration data from the NanoDrop spectrophotometer, DEPC-treated H_2O was added to 0.2ml micro-centrifuge tubes such that a final volume of 13µl including primers and RNA was achieved. 1µl of 10mM dT-adapter primer and 2µg of RNA was added. The contents of the tubes were mixed, and incubated for

seven minutes at 70°C to denature the template RNA before chilling on ice. The tubes were centrifuged briefly to collect the contents at the bottom of the tube. A master mix containing 4µl of 5× first strand buffer, 1µl of 0.1M DTT, 1µl 10mM deoxynucleotide (dNTP) mix, and 1µl of Superscript^{TM} III reverse transcriptase (Invitrogen^{TM}) was added to each tube and mixed gently. The tubes were centrifuged briefly to collect the contents at the bottom of the tube, and were incubated in a PCR machine for five minutes at 22°C, two hours at 50°C for complementary DNA (cDNA) synthesis, and 15 minutes at 70°C to terminate the reverse transcriptase, before being promptly removed and chilled on ice. The cDNA was diluted 1:10 with DEPC-treated H_2O, and stored at -20°C until required.

2.9.4. Quantitative PCR

Quantitative PCR was used to quantify the gene expression of wheat grain samples collected during this study, and was performed on cDNA synthesised from RNA extracted from developing grain. The full tehnical protocol is presented in appendix A.7.

A master mix was prepared containing 0.7µl of 10µM sense primer, 0.7µl of 10µM anti-sense primer, 0.03µl of 100× ROX internal reference dye, 11.22µl of H_2O, and 13.75µl of SYBR® Green JumpStart^{TM} Taq ReadyMix^{TM} (Sigma Aldrich, Gillingham, UK) per sample. 26.4µl of master mix and 1.1µl of cDNA was added to a 1.5ml micro-centrifuge tube, mixed well, and centrifuged to collect the contents. 25.1µl was loaded into each well of a white 96-well PCR plate, which was sealed with a transparent qPCR seal, and centrifuged to collect the contents. The completed plate was loaded into an Applied Biosystems 7500 qPCR machine (Applied Biosystems, Foster City, CA, USA) and run for two minutes at 50°C, ten minutes at 95°C, 41 cycles of 15 seconds at 95°C and one minute at 60°C, 15 seconds at 95°C, and a final 15 seconds at 60°C.

2.9.5. Analysis of quantitative PCR data

The results from qPCR were analysed using the Applied Biosystems 7500 software (v2.05). Rn values were exported, and mean primer efficiency values calculated using the LinRegPCR software (v12.3, Ruijter *et al.* (2009)). Analysis of the Ct values with relation to the primer efficiency values was conducted using the normalised relative quantification (NRQ) method (Rieu *et al.* 2009), which relies on the expression of the 18S and 28S rRNA reference genes for normalisation. For calculating the NRQ values, the following formula was used:

$$\text{NRQ} = \frac{\text{Target primer efficiency}^{-\text{ Target Ct}}}{\sqrt{(\text{18s primer efficiency}^{-\text{ 18s Ct}})\times(\text{28s primer efficiency}^{-\text{ 28s Ct}})}}$$

Log transformed ($\log_2(\frac{1}{\text{NRQ}})$) NRQ values were then used for statistical analysis to describe the relative expression of protein synthesis genes during grain-filling.

2.10. **Statistical analysis**

The GenStat® statistical software package (2015, Eighteenth Edition, VSN International Ltd, Hemel Hempstead, UK) was used to analyse all data collected during this study.

Analysis of variance (ANOVA) was applied to all single-variate data, including all morpholog-

ical data (yield, thousand grain weight, grain count etc.), grain nitrogen concentration data, and SE-HPLC data. The least significant difference (LSD) at the 5% (P=0.05) level, calculated from the standard error of the difference (SED) between means on the residual degrees of freedom (DF) from the ANOVA was used to make comparison of relevant means.

Restricted maximum likelihood (REML) was used to fit linear mixed models in instances where the application of ANOVA was unsuitable or inadequate, namely in the analysis of the SDS-PAGE protein composition data, the light-microscopy protein gradient and protein body size-distribution data, and the RNA expression data. In the case of the analysis of the SDS-PAGE data, REML was used in order to account for the complex design structure required to account for variation across protein gels, and between gel electrophoresis tanks (which each contained two gels). For the analysis of light-microscopy and RNA expression data, REML was used to regress on certain variables in the model, which would not be possible in ANOVA. With regards to the analysis of the light-microscopy data, this allowed for the variable associated with distance from the aleurone layer to be included as a factor in the analysis. This approach was essential for the analysis of the continuous size-distribution data, and beneficial in the analysis of the discrete protein concentration gradient data, which consisted of protein concentration measurements attributed to one of five endosperm zones. These zones were converted to physical distance (which varied between different treatment combinations), and this variable used as a factor in the analysis to provide further insight to this data. As a result, any difference in the size of grain between treatments was accounted for in the analysis of the data. A similar approach was taken in the analysis of the RNA expression data, with the accumulated thermal time variable included as a factor in the model, allowing for differences in the experienced thermal time to be accounted for.

Canonical variate (CV) analysis was used in the exploratory analysis of SDS-PAGE protein composition data to illustrate the effect, and magnitude of effect, that each treatment combination had on protein composition. Individual (in the case of the controlled-environment experiment) and grouped (for the WGIN field experiment) protein measurements were used as the terms in the CV analysis, with a grouping factor relating to the specific treatment combination. In the controlled-environment experiment, this grouping provided separation of protein composition data by each combination of temperature (20/28°C) and nitrogen (low/high) treatment over the six sampling timepoints; and in the WGIN diversity field experiment, the grouping consisted of each combination of year (2015–2017) and nitrogen treatment (100kg-N/ha/350kg-N/ha) for each of the four genotypes used in the study. The loadings of each protein/protein group are also presented, as an indication of which protein or group of proteins were most affected by the applied treatments. It should be noted that whilst CV analysis generally requires higher levels of replication that available in this study, this does not preclude its application for investigative purposes when used in combination with REML analysis.

Simple linear regression was used for the comparison of storage protein synthesis gene expression data with the SE-HPLC measurements of the relevant protein. The NRQ values were summed over all sampling timepoints to produce a value for accumulated NRQ, and this value was regressed against the absolute measurements of the appropriate SE-HPLC protein fraction. The abundance of the F1 fraction was regressed against the relative expression of the HMW-glutenin transcript, the F2 fraction against the expression of the LMW-glutenin transcript, the F3 fraction against the ω-gliadin transcript, and the F4 fraction was regressed against the sum of the expression of the α- and β-gliadin and the mean of the two γ-gliadin transcripts. The relevant coefficient of determination (R^2) and P values are presented for comparisons between gene expression, and abundance of the corresponding protein.

Chapter 3: Meteorological data

3.1. Introduction

Meteorological data was recorded from the on-site weather station at Rothamsted Research, Harpenden, UK, and provided daily measurements for mean temperature, rainfall, and hours of sun for the duration of the WGIN field trial experiments held in 2015–2017. From this data, accumulated thermal time was calculated to record how developmental speed and timing may vary between years. Only the data for the period of sampling during grain-filling is reported (0–35DPA), since this covers time period of interest to this project, and extends to the late stages of grain-filling. For figures 3.1, 3.2, 3.3, and 3.4, the period of time covering the 35-day sample period for Cadenza is shown.

3.2. Temperature and accumulated thermal time

Mean temperature data for the first 35 days of grain-filling in Cadenza for 2015–2017 is presented in figure 3.1. Events of note include prolonged periods of high temperatures from 20–26 days after anthesis in 2015, and between 14–23 days after anthesis in 2017. The period of high temperature in 2017 is of particular interest as it falls within a critical time for grain-filling with regards to protein accumulation. The year of 2016 was relatively unremarkable in comparison, with no prolonged periods of hot weather, and stable temperatures day-to-day.

From the mean daily temperature measurements accumulated thermal time was calculated, and is presented in figure 3.2. The accumulated thermal time data show a linear increase in accumulated thermal time in 2016 due to the relatively stable mean temperatures experienced. However, the periods of elevated temperature are evident in the data from 2015 and 2017, with 2017 in particular showing a rapid increase in accumulated thermal time in mid grain-filling. The total accumulated thermal time over the 35 day period covering grain-filling was comparable between 2015 and 2016, but was greater in 2017.

3.3. Rainfall and sunlight

Accumulated hours of sunlight was measured for the duration of grain-filling, and the data from the first 35 days of grain-filling in Cadenza is presented in figure 3.3. The years of 2015 and 2017 showed comparable levels of sunlight by the end of this 35-day period, with increased sunlight that mirror the daily temperature measurements shown in figure 3.1. There was less sunlight throughout grain-filling, and less accumulated sunlight by 35DPA in 2016.

Rainfall during grain-filing is presented in figure 3.4 as total accumulated rain. This data shows that 2017 was a particularly dry year during grain-filling, with approximately half of the accumulated rainfall of 2016, and with a prolonged spell over approximately two weeks with no recorded rainfall. The situation was similar in 2015, although rainfall was more regular than 2017, without any prolonged period with no rain.

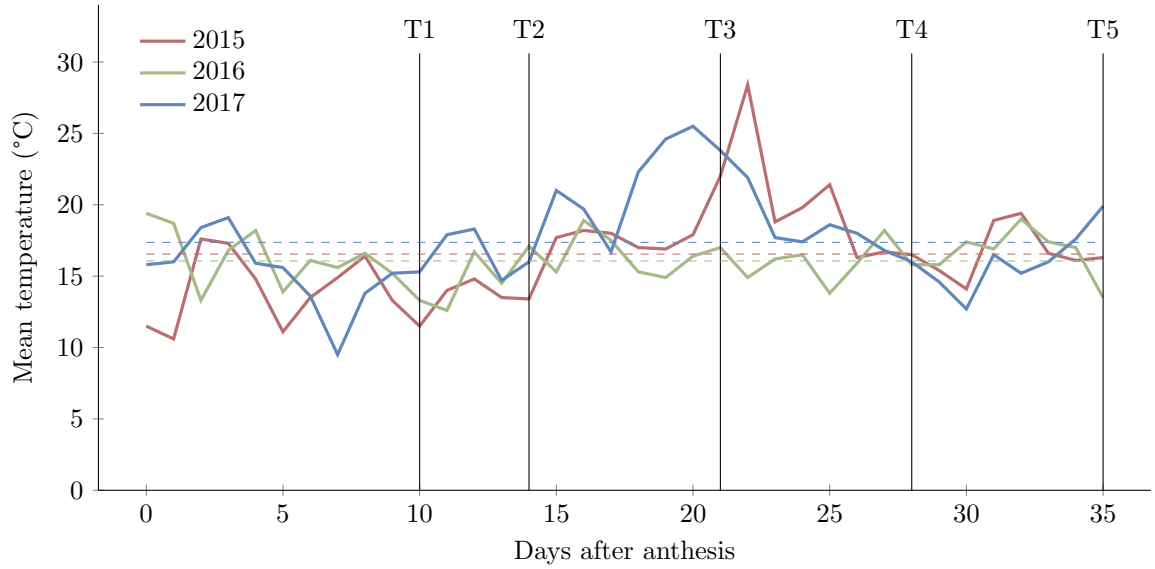


Figure 3.1: **Average temperature during grain-filling for the WGIN diversity experiment.** Mean daily temperatures are presented for the 35 days of grain-filling following the anthesis of Cadenza, over the three years (2015–2017) of the WGIN diversity field experiment, with overall mean shown as a dashed line. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

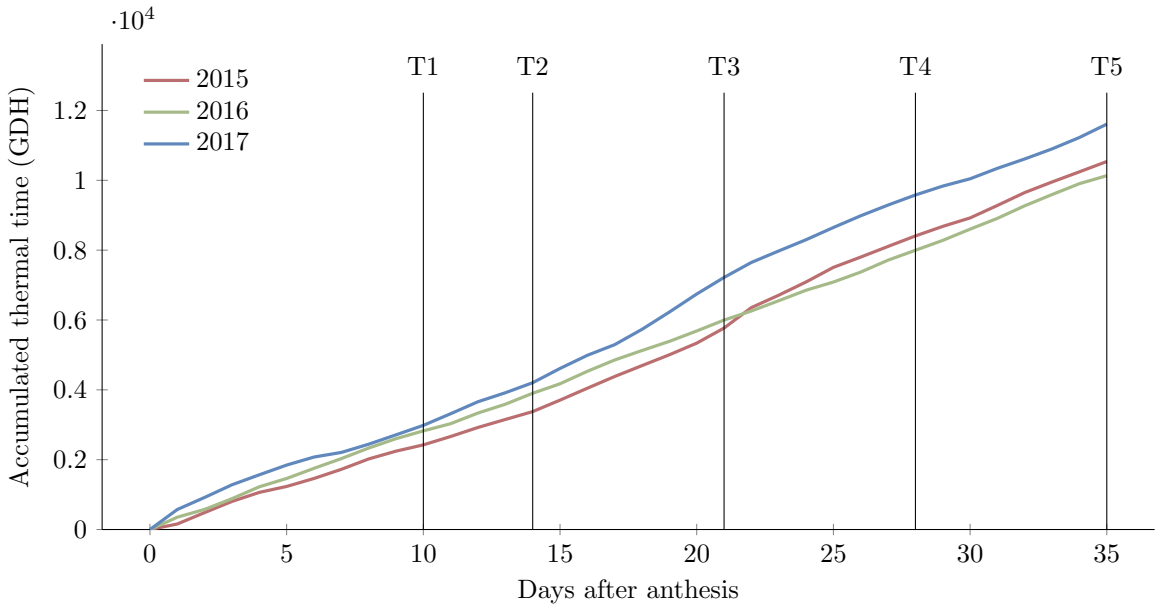


Figure 3.2: **Accumulated thermal time during grain-filling for the WGIN diversity experiment.** Accumulated thermal time was calculated from the daily mean temperature using a base temperature of 4.1°C, and expressed as GDH. Data is presented for the 35 days of grain-filling following the anthesis of Cadenza, over the three years (2015–2017) of the WGIN diversity experiment. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

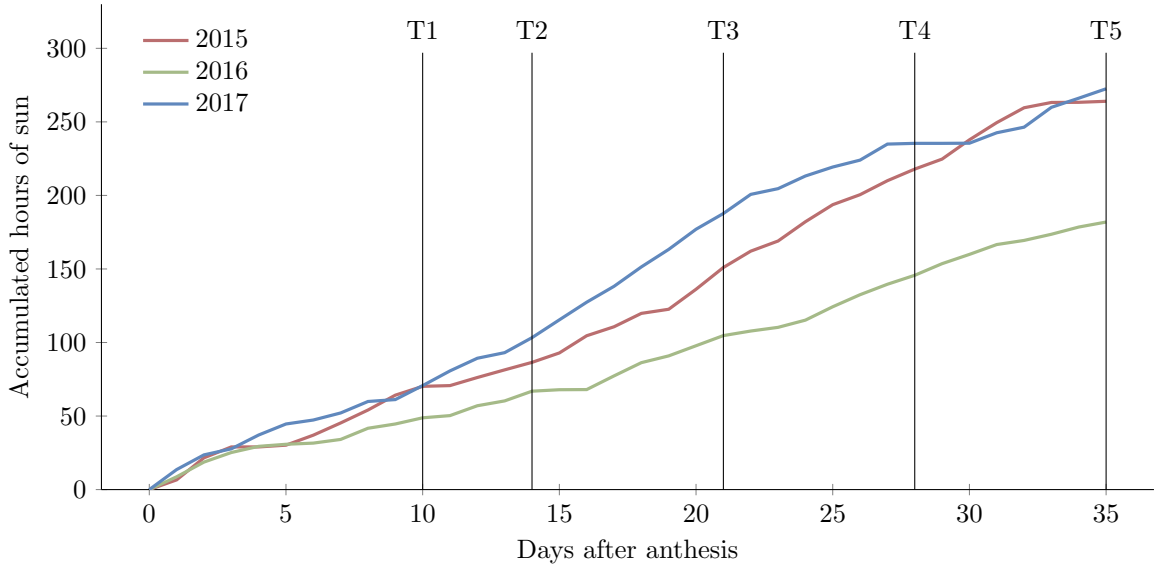


Figure 3.3: **Radiation accumulated during grain-filling for the WGIN diversity experiment.** Hours of sun data is presented for the 35 days of grain-filling following the anthesis of Cadenza, over the three years (2015–2017) of the WGIN diversity experiment. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

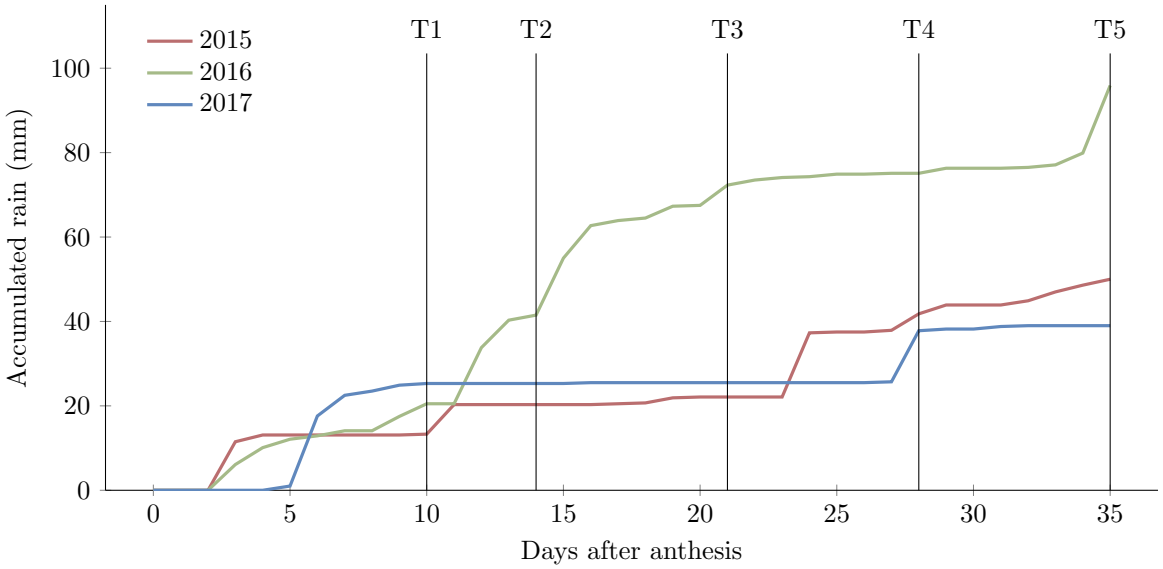


Figure 3.4: **Accumulated rainfall during grain-filling for the WGIN diversity experiment.** Data is presented for the 35 days of grain-filling following the anthesis of Cadenza, over the three years (2015–2017) of the WGIN diversity experiment. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

Table 3.1: Anthesis dates of the four wheat varieties sampled in the WGIN diversity field trial experiment. Cadenza, Hereward, and Istabraq show similar antehsis dates, whilst the photoperiod-insensitive Soissons consistently reached anthesis earlier.

Variety	Anthesis in:	2015	2016	2017
Cadenza		09/06	07/06	30/05
Hereward		10/06	09/06	31/05
Istabraq		10/06	09/06	31/05
Soissons		03/06	30/05	25/05

3.4. **Comparisons of weather experienced by different genotypes**

Of the four different genotypes sampled in this study, Soissons was notable for its early flowering. Whilst Cadenza, Hereward, and Istabraq all reached anthesis within one or two days of each other, Soissons flowered between five to ten days earlier, as shown in table 3.1. Due to this, the climate experienced by Soissons was different to that experienced by the other varieties. The accumulated thermal time, hours of sun, and rainfall were compared between Cadenza and Soissons in each of the three years sampled as part of this study. Only these two genotypes are compared for clarity, as differences between the conditions experienced between Cadenza, Hereward, and Istabraq were minimal.

Comparisons of the accumulated thermal time experienced by Cadenza and Soissons during the first 35 days of grain-filling are presented in figure 3.5. These comparisons show how the prolonged periods of hot weather shown in figure 3.1 occurred later in development for Soissons. Furthermore, they show how the accumulated thermal time by 35 days was comparable in 2015 and 2017, but was lower for Soissons in 2016 due to cooler weather in the early days of grain-filling.

Figure 3.6 shows how hours of sunlight varied between Cadenza and Soissons for the first 35 days of grain-filling for. In 2015, Soissons had more sun over the first 21 days of grain-filling, and ultimately accumulated more hours of sun over the duration of sampling. Differences between the two genotypes were less pronounced in 2016 and 2017, with similar levels of sun throughout grain-filling.

The differences in rainfall during grain-filling for Cadenza and Soissons are presented in figure 3.7. In 2015 and 2016, Soissons generally saw less rainfall that Cadenza at the same time after anthesis. The dry periods in 2015 and 2017 discussed in section 3.3 occurred later in development for Soissons, which may be favourable, since water scarcity would occur closer towards the dehydration stage of grain-filling.

3.5. **Conclusions and discussion**

The meteorological data collected as part of the WGIN diversity field trial experiment revealed some clear differences between the three years sampled as part of this study. With regards to temperature and accumulated thermal time, the years of 2016 and 2017 stand out: 2016

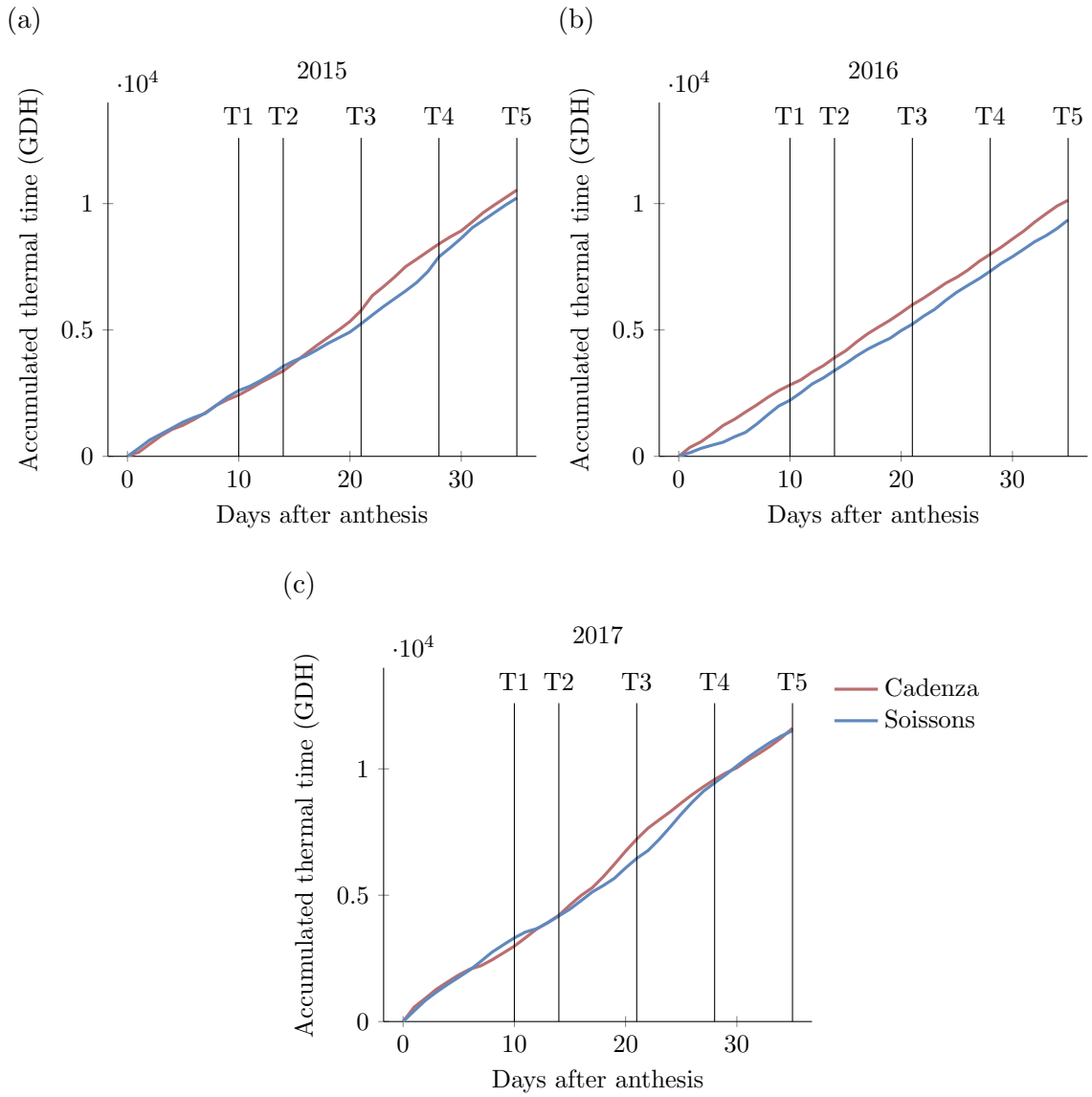


Figure 3.5: **Comparison of the accumulated thermal time for Cadenza and Soissons during grain-filling in the WGIN diversity experiment.** The differences in accumulated thermal time in the first 35 days of grain-filling following anthesis for Cadenza and the early-flowering Soissons are presented. Data is presented for all three years of the WGIN diversity experiment: (a) 2015, (b) 2016, and (c) 2017. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

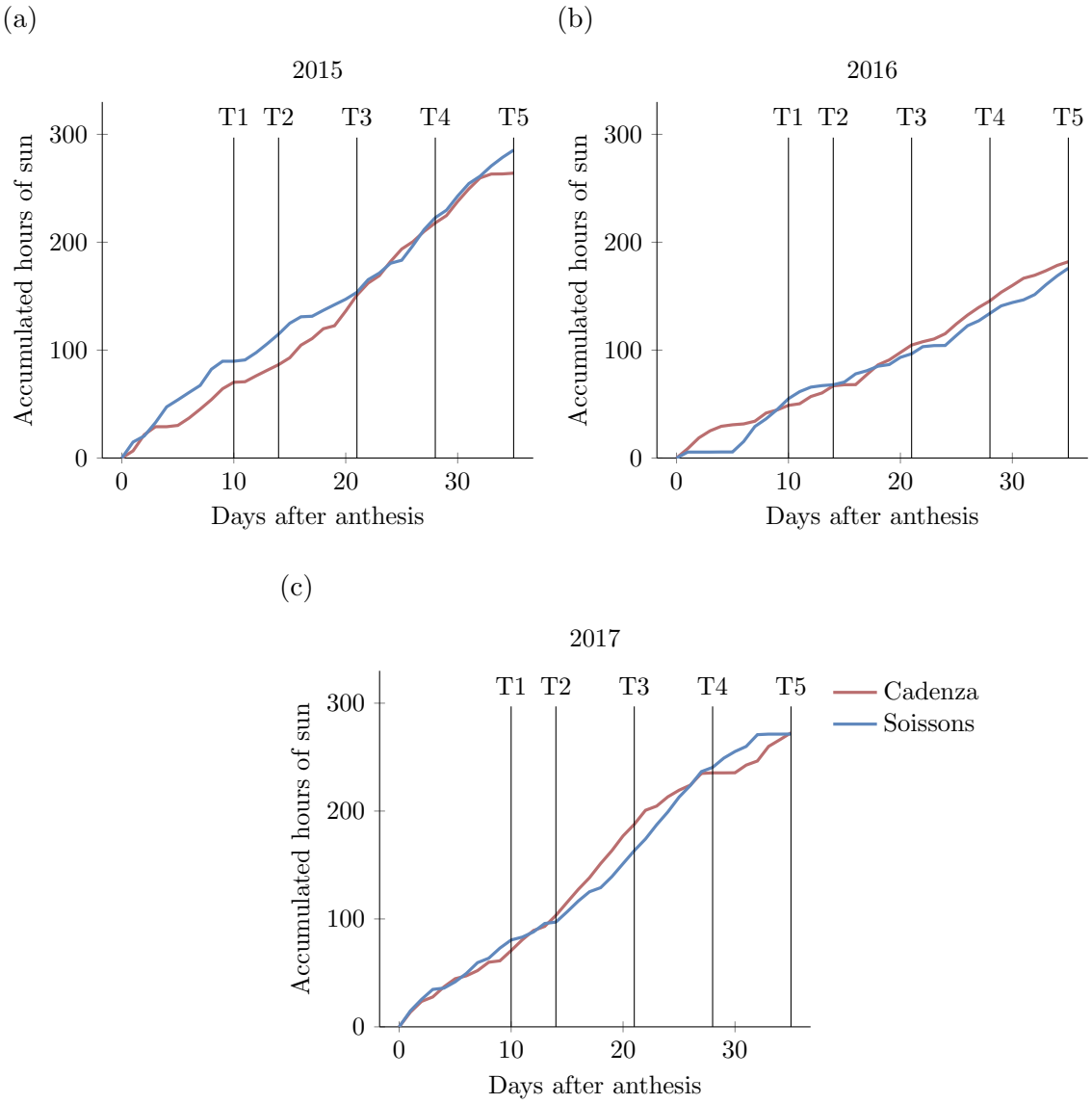


Figure 3.6: **Comparison of the accumulated radiation for Cadenza and Soissons during grain-filling in the WGIN diversity experiment.** The differences in radiation (presented as accumulated hours of sun) in the first 35 days of grain-filling following anthesis for Cadenza and the early-flowering Soissons are presented. Data is presented for all three years of the WGIN diversity experiment: (a) 2015, (b) 2016, and (c) 2017. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

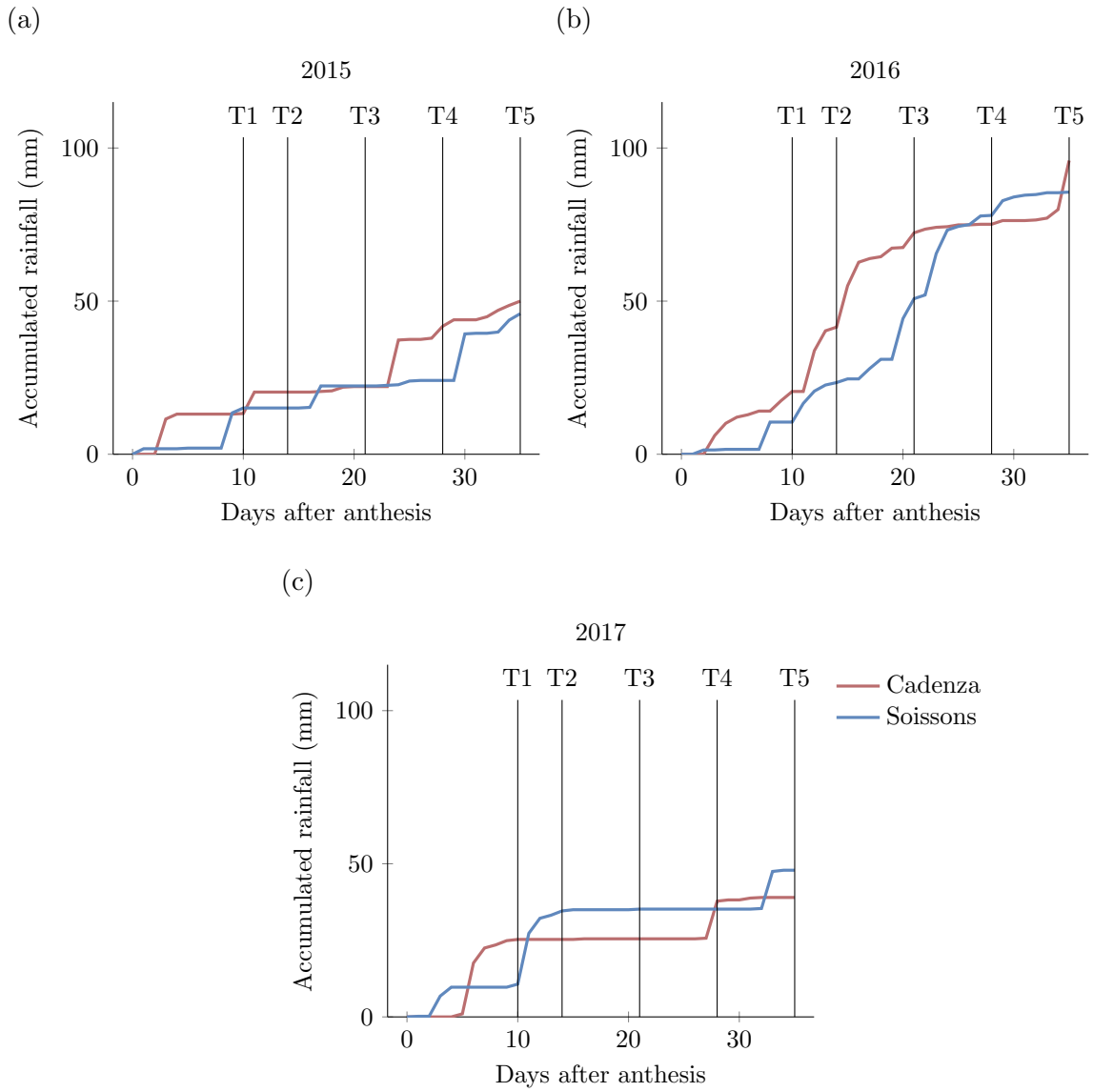


Figure 3.7: **Comparison of the rainfall for Cadenza and Soissons during grain-filling in the WGIN diversity experiment.** The differences in rainfall in the first 35 days of grain-filling following anthesis for Cadenza and the early-flowering Soissons are presented. Data is presented for all three years of the WGIN diversity experiment: (a) 2015, (b) 2016, and (c) 2017. Vertical lines represent the five timepoints (T1–T5) at which grain were sampled for analysis.

was a mild year, with stable temperatures throughout the grain-filling period, whilst in 2017 there was a nine day period of elevated temperature during early to mid grain-filling. There was a shorter heatwave in 2015, slightly later in grain-filling, which likely had a lesser effect on the development of the plants than the heatwave in 2017. The amount of thermal time accumulated throughout grain-filling was greatest in 2017, driven by both the heatwave and higher temperatures throughout grain-filling, and was lowest in 2016. The short heatwave in 2015 resulted in a slight boost in accumulated thermal time compared to 2016, but ultimately 2015 and 2016 accumulated similar amounts of thermal time. From the temperature data collected, we could predict that 2015 and 2016 would achieve higher yields due to the slower accumulation of thermal time, whilst the high temperatures of 2017 would reduce yield, mainly by reducing starch content, thereby increasing the concentration of nitrogen within the grain.

In addition to temperature data, the collection of sunlight and rainfall data also showed some differences between the three years. Both 2015 and 2017 saw approximately 150% of accumulated sunlight during grain-filling compared to 2016, with increases in accumulated sunlight correlating with the heatwaves in 2015 and 2017. The lower accumulated sunlight in 2016 was reflected in the rainfall data: 2016 saw approximately 200% of the rainfall of both 2015 and 2017 during grain-filling, with no period without rain of more than three days during this period. The driest year was 2017, which also saw a period of 18 days with no rain, which spanned from early to late grain-filling in Cadenza. 2015 had more consistent rainfall, and accumulated more than in 2017, but was still considerably drier than 2016.

In summary, 2015 saw relatively mild temperatures during grain-filling, with a short (five days) heatwave at mid to late grain-filling, and a low level of accumulated thermal time when compared to 2017. The mild temperatures in 2015 were also accompanied by high levels of sunlight, and relatively consistent rainfall without a prolonged period with no rain. 2016 was the most consistent year temperature-wise, with mild temperatures throughout grain-filling. However, 2016 saw less sunlight than the other years, as well as significantly more rain. 2017 was particularly noteworthy due to the occurrence of a prolonged heatwave, as well as generally higher temperatures, which resulted in a much greater amount of accumulated thermal time during grain-filling. Sunlight was high during 2017, but rainfall was the lowest of the three years, with a prolonged period (18 days) with minimal rain that lasted from early to late grain-filling. Of the three years, it could be predicted that 2015 would show the best yield and grain quality due to the mild temperatures, high level of sunlight, and regular rain. The differences between 2016 and 2017 are perhaps too great to make a prediction on which would produce the highest yield. Whilst 2017 saw the highest temperatures, and the least rainfall, it also experienced a much greater amount of sunlight during grain-filling than 2016. Since in UK agriculture sunlight is more likely to be limiting yield than rainfall (AHDB 2018), it is possible that the increased rainfall and mild temperatures in 2016 may be counteracted by the lack of sunlight with regards to yield.

Of the three years of the WGIN field trial experiment, 2016 and 2017 stand out as the most contrasting years, with 2016 representing a consistently mild year with adequate rainfall, and 2017 a year with a heatwave and prolonged period with no rain during grain-filling. Since one of the aims of this study is to investigate the effect that climate change, and increased frequency of heatwave and drought, may have on grain-filling, the comparison of data between 2016 and 2017 is of great interest. This data was used to inform the decision of which two years should be compared with RNA expression analysis; the years of 2016 and 2017 were chosen to increase the chances of identifying a year-on-year difference in gluten storage protein synthesis gene expression.

The photoperiod-insensitive variety Soissons was included in the analysis to investigate the effect that early anthesis might have on grain yield and quality. When compared to the photoperiod-sensitive varieties, Soissons flowered five to ten days early, resulting in considerable differences in the timing of climate events with regards to grain development. Analysis of the temperature data collected during grain-filling showed that Soissons accumulated less thermal time in 2015 and 2016 than the other genotypes, and a comparable amount in 2017. The differences in accumulated thermal time were greatest in 2016, where Soissons experienced particularly low temperatures during the first six days of grain-filling. From the information presented in figure 3.5, it is clear that the accumulated thermal time at a particular sampling timepoint (T1–T5) was often different between varieties, and so it is likely that development was not at a directly comparable stage between the photoperiod-sensitive and -insensitive varieties. Again, this difference is particularly evident in 2016.

As with temperature, there were differences in the amount of sunlight and rainfall experienced by Soissons when compared to the photoperiod-sensitive varieties. In 2015, Soissons experienced more sunlight during early grain-filling, and accumulated more hours of sun by 35DPA, in 2016 Soissons generally saw less sun throughout grain-filling, and accumulated slightly fewer hours of sun in total than the other varieties, and in 2017 Soissons accumulated radiation more slowly during mid grain-filling, but ultimately accumulated the same amount of sunlight over the first 35 days of grain-filling as the other varieties. The differences in rainfall between the different varieties are likely less important than the differences in sunlight and temperature, since the soil's ability to hold water buffers the plants against periods without rain to a certain degree. Nevertheless, Soissons experienced less total rainfall in the first 35 days of grain-filling in 2015 and 2016, but had more rain in 2017. However, since rainfall was generally plentiful, perhaps the only considerable difference was the timing of the period of drought in 2017, with the drought occurring earlier in grain-filling for Soissons compared to the other genotypes.

Whilst the differences in yield and grain quality between the early-flowering Soissons and the other varieties may be obscured by any number of factors unrelated to climate, it may be possible to identify climactic effects by looking at the differences between years and genotypes.

Unfortunately since no final harvest data for Soissons was available in 2017, only 2015 and 2016 can be used for comparison of factors such as yield and protein composition. However, measurements for nitrogen content were taken throughout grain-filling for all three years, and as such it may be possible to analyse the effect that the timing of the heatwave and drought in 2017 had on the protein accumulation in each cultivar. If an increase in nitrogen concentration, relative to other years, was observed during the 2017 heatwave in the three late-flowering varieties but not Soissons, it could be speculated that the later timing of the increased temperature reduced the severity of the heatwave on nitrogen concentration in the grain. Aside from the dramatic differences between 2017 and 2016, it may also be possible to identify the effect that some of the less pronounced yearly differences had on the early-maturing Soissons. For instance, the difference in accumulated thermal time between Soissons and the other cultivars was less in 2015 than in 2016. This difference could manifest itself as Soissons showing higher yields in 2016, relative to the other cultivars.

Chapter 4: Grain yield and yield component results

4.1. Introduction

With a growing population, maximising the yield of our crops is a shared aim between farmers, breeders, and agricultural scientists. In wheat, yield is determined by both the size and number of grain produced. Whilst the weight of an individual grain is simple to characterise, determining and understanding the grainset of the mature wheat plant is more complicated. Tiller and ear number, spike size, and fertility of individual spikelets all interact to determine the number of grain harvested from any one plant. Both grain yield components, size and number, are sensitive to changes in temperature (Spiertz *et al.* 2006; Rahman *et al.* 2009), solar radiation (Fischer 1985), nitrogen regime (Otteson *et al.* 2007), and genotype (Slafer *et al.* 2014), and as such understanding the response wheat has to each of these factors is of great importance. To this aim, two experiments were completed as part of this study: a controlled-environment study looking at the combined effects of nitrogen input and elevated temperature during grain-filling, and the WGIN diversity field trial, from which four wheat genotypes grown under either low (100kg-N/ha) or high (350kg-N/ha) nitrogen regimes were sampled over three consecutive years.

4.2. Controlled-environment experiment

In the controlled environment experiment, UK spring wheat cultivar Cadenza was grown to anthesis under either low or high nitrogen input, before being exposed to either a control (20°C) or high (28°C) daytime temperature treatment for the duration of grain-filling. At harvest measurements were taken for total grain yield and thousand grain weight. From this data grain count was calculated, allowing for any changes in yield to be attributed to changes in grain weight, changes in grainset, or a combination of these factors. Additionally, dimensional measurements for grain area, length, and width were taken from a sub-sample of mature grain to describe how individual grain morphology is affected by temperature and nitrogen. Finally, to quantify the effect that the two nitrogen treatments used had to the nitrogen status of the plants, leaf chlorophyll content measurements were taken using a SPAD meter.

4.2.1. Yield

Total grain yield was measured on grain harvested from plants at full maturity, bulking the grain from all five plants within each pot before oven drying to 5% moisture.

Grain yield was decreased by elevated temperature during grain-filling, and increased by the high nitrogen treatment applied prior to anthesis, with a significant interaction between these factors ($F_{1,6}$=15.28, P=0.008, SED=0.324 on 7.34 DF) (figure 4.1). Of these two factors, temperature had the greatest effect, with an average decrease in yield of 34% when elevated temperatures were experienced during grain-filling. The effect of nitrogen was overall smaller

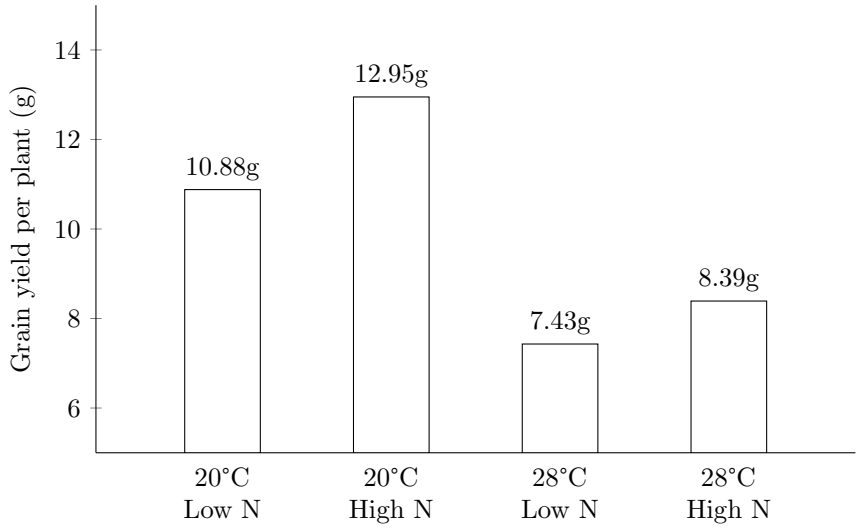


Figure 4.1: **Grain yield is increased by high nitrogen input and elevated post-anthesis temperature.** Mean grain yield from the controlled-environment experiment, presented as individual treatment combinations of control temperature (20°C), low- and high-nitrogen, and elevated temperature (28°C), low- and high-nitrogen. LSD (at the 5% level) of 0.49 for comparing means within the same level of temperature treatment, and 0.76 for all other comparisons.

.

than the effect of temperature, but was also different depending on the temperature treatment, with increased nitrogen supply resulted in a greater yield increase (19%) under the control temperature treatment than under the elevated temperature treatment (13%).

4.2.2. Thousand grain weight

Thousand grain weight was measured on a sub-sample of the grain used to measure yield, *i.e.* grain bulked from five plants within the same pot at maturity. As a representative measure of individual grain weight, thousand grain weight (TGW) can help to describe the mechanisms behind changes in yield, and can indicate whether any increases in yield are due to increases in the weight of individual grain, or increases in the number of grain produces. The mean thousand grain weights for each treatment combination are presented in full in table 4.1.

Thousand grain weight was significantly decreased by elevated temperature during grain-filling ($F_{1,5}$=67.47, P<0.001, SED=1.456 on 5 DF) (figure 4.2a), with an average decrease of 30%

Table 4.1: Full thousand grain weight dataset from the controlled-environment experiment presented per treatment combination, and averaged over experimental blocks.

Treatment	Thousand grain weight (g)
20°C, low-nitrogen	38.52
20°C, high-nitrogen	41.59
28°C, low-nitrogen	27.56
28°C, high-nitrogen	28.62

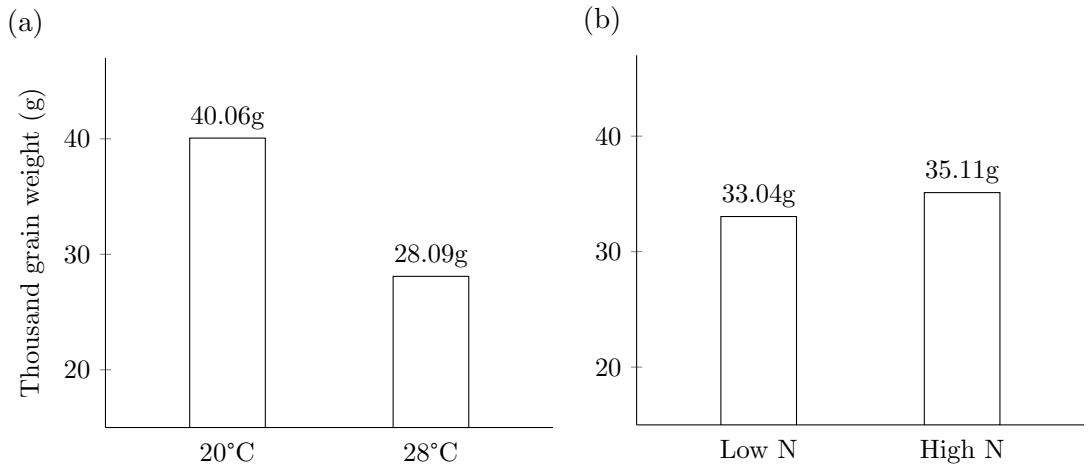


Figure 4.2: **Thousand grain weight is reduced by high temperatures during grain-filling, and increased by nitrogen fertilisation.** Mean TGW from the controlled-environment experiment, showing the effects of (a) post-anthesis temperature, LSD (at the 5% level) of 3.74, and (b) nitrogen supply during vegetative development, LSD of 1.59.

.

between the control- and high-temperature treatments used in this experiment. Limiting the nitrogen supply prior to anthesis also resulted in a decrease in TGW ($F_{1,6}$=10.19, P=0.019, SED=0.648 on 6 DF) (figure 4.2b), with TGW 6% higher, on average, when nitrogen supply was increased. There was no significant interaction between temperature and nitrogen treatments ($F_{1,6}$=2.50, P=0.165).

4.2.3. Total grain count

To complete the description of the factors contributing to grain yield, a total grain count was calculated from the yield and TGW data. In combination with TGW, this information can tell us whether changes in yield were due to increases in grain size, increases in grain number, or a combination of these two factors. Calculated grain counts for each treatment combination are presented in table 4.2.

Limiting nitrogen supply to the plants during vegetative development resulted in a 10% decrease in the total grain count ($F_{1,6}$=26.17, P=0.002, SED=5.35 on 6 DF) (figure 4.3).

Table 4.2: Full grain count dataset from the controlled-environment experiment. Grain count calculated from yield and TGW data, and is presented per treatment combination, averaged over experimental blocks.

Treatment	Grain count per plant
20°C, low-nitrogen	281
20°C, high-nitrogen	314
28°C, low-nitrogen	270
28°C, high-nitrogen	296

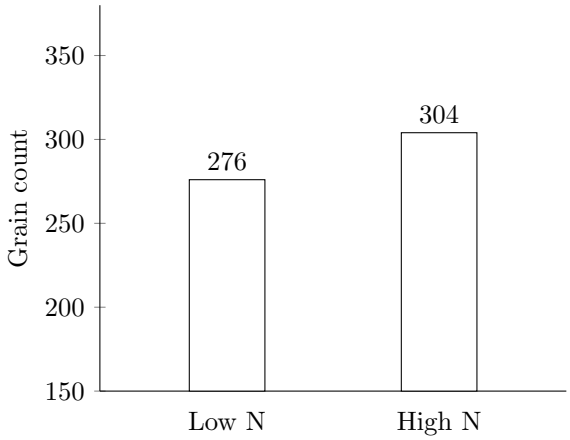


Figure 4.3: **Grain count is increased by nitrogen fertilisation.** Calculated grain count data from the controlled-environment experiment, grouped by nitrogen treatment and averaged over temperature treatments. LSD (at the 5% level) of 13.

However, the elevated post-anthesis temperature treatment had no significant effect on grain count ($F_{1,5}$=2.20, P=0.198), with an average of 297 grain per plant under the control treatment, and 283 when the high-temperature treatment was applied (LSD at the 5% level of 25). Likewise, no evidence was found of a significant interaction between temperature and nitrogen treatments ($F_{1,6}$=0.14, P=0.720).

4.2.4. Grain dimension measurements

Grain area, length, and width was measured for grain sampled at maturity. These measurements provide data for grain size, which is complementary to the TGW data, and describes how different dimensions of the grain may be affected by nitrogen supply, and post-anthesis temperature. The grain size analysis measurements are presented in full in table 4.3.

Elevated post-anthesis temperature decreased grain area by an average of 13%, from 19.12mm^2 to 16.63mm^2 ($F_{1,4}$=28.43, P=0.006, SED=0.330 on 4 DF) (figure 4.4a). Nitrogen supply had a smaller effect on grain area, with the low-nitrogen treatment reducing grain area by 3% from 17.58mm^2 to 18.17mm^2 ($F_{1,22}$=7.48, P=0.012, SED=0.213 on 22 DF) (figure 4.4b). No significant interaction was found between temperature and nitrogen treatments on grain area ($F_{1,22}$=0.26, P=0.617).

Grain length and width were differentially affected by the temperature and nitrogen treatments used in this experiment: grain length was increased by 3% under the high nitrogen treatment ($F_{1,22}$=18.1, P<0.001, SED=0.0451 on 22 DF) (figure 4.4c), whilst grain width was reduced by 12% under the elevated post-anthesis temperature treatment ($F_{1,4}$=96.57, P<0.001, SED=0.0441 on 4 DF) (figure 4.4d). Temperature did not have a significant effect on grain length ($F_{1,4}$=0.09, P=0.776), and nitrogen did not have a significant effect on grain width ($F_{1,22}$=0.78, P=0.387). Likewise, no interactions were found between these factors for either grain length ($F_{1,22}$=0.55, P=0.467), nor width ($F_{1,22}$=0.09, P=0.771).

Table 4.3: Full dataset for grain area, length, and width measurements of mature grain from the controlled-environment experiment, presented by treatment combination and averaged over experimental blocks. Measurements were made on sub-samples of 100 grain using the automated MARVIN grain analyser.

Treatment	Area (mm^2)	Length (mm)	Width (mm)
20°C, low-nitrogen	18.88	6.28	3.72
20°C, high-nitrogen	19.36	6.44	3.73
28°C, low-nitrogen	16.28	6.29	3.28
28°C, high-nitrogen	16.98	6.52	3.31

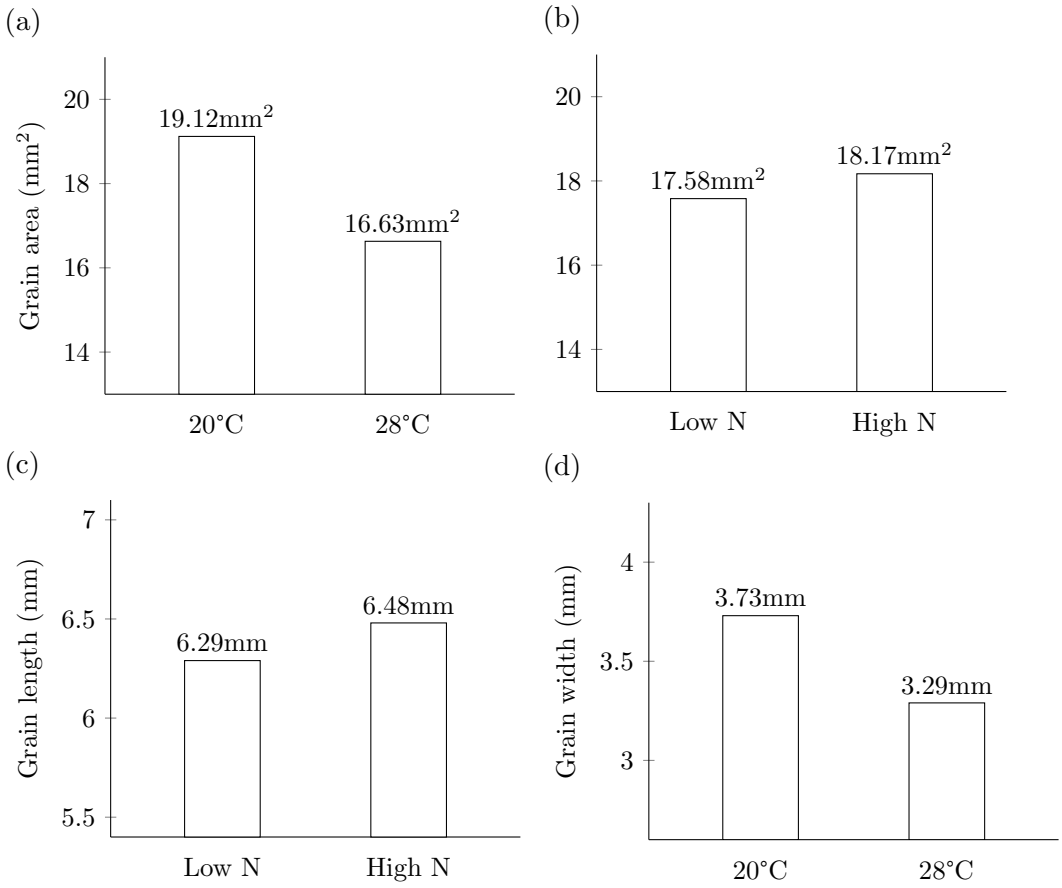


Figure 4.4: **Elevated post-anthesis temperature reduces grain area and width, whilst nitrogen fertilisation increases grain area and length.** Grain dimension measurements from the controlled-environment experiment grouped by significant effects: (a) mean grain area for each temperature treatment, LSD (at the 5% level) of 1.30, and (b) for each nitrogen treatment, LSD of 0.44; (c) mean grain length for each nitrogen treatment, LSD of 0.09; (d) mean grain width for each temperature treatment, LSD of 0.12.

4.2.5. SPAD measurements

SPAD measurements were taken to identify differences in leaf nitrogen content, primarily to assess the impact of the two different nitrogen treatments on the nitrogen status of the plants. Measurements were taken once all nutrient solutions had been applied, five days after anthesis was complete for all plants sampled. The aim of this was to identify the effect that the two nitrogen treatments had on the plants, since it was impossible to visually detect a difference between the two treatments used. Full results are shown in table 4.4.

The low-nitrogen treatment reduced the leaf chlorophyll content measured by SPAD ($F_{1,16}$=18.04, P<0.001, SED=0.385 on 16 DF), decreasing the average measurement from 53.6 to 51.9 (with an LSD of 0.8 at the 5% level) (figure 4.5). The temperature treatment, which had only been applied for approximately five days, had no significant effect on the leaf chlorophyll content ($F_{1,16}$=0.23, P=0.639), and there was no interaction between nitrogen and temperature treatments found ($F_{1,16}$=1.93, P=0.183).

Table 4.4: Full dataset of leaf chlorophyll content measured using a SPAD meter from the controlled-environment experiment. Measurements were taken five days after anthesis from three points on the flag leaves of all five plants in a pot, with eight pots sampled for each treatment combination. Presented figures are average values for each combination of temperature and nitrogen treatment.

Treatment	SPAD measurement
20°C, low-nitrogen	52.10
20°C, high-nitrogen	53.20
28°C, low-nitrogen	51.75
28°C, high-nitrogen	53.92

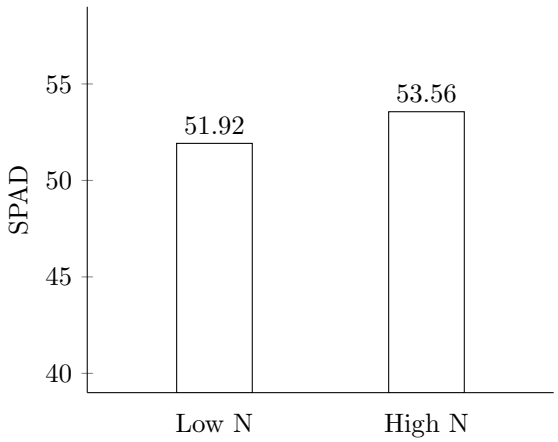


Figure 4.5: **Leaf chlorophyll content is increased by nitrogen fertilisation.** SPAD measurements taken after anthesis from the controlled-environment experiment, presented by nitrogen treatment and averaged across temperature treatments. LSD (at the 5% level) of 0.82.

.

4.3. **WGIN diversity field experiment**

Over three years of the WGIN diversity field trial, grain yield, TGW, number of ears, and grain per ear was measured for four wheat genotypes: Cadenza, the group 2 bread-making wheat used in the controlled-environment experiment; Hereward, a group 1 bread-making wheat; Istabraq, a group 4 feed wheat; and Soissons, an early-flowering group 2 bread-making wheat. This range of genotypes was sampled to capture differences in the responses shown to both nitrogen input and to the climate experienced during grain-filling, which varied from year to year, and is discussed in detail in chapter 3.

Measurements for yield and TGW were made on grain machine-harvested at maturity from the 3m × 9m field plots after bench drying to 15% moisture, ear counts were made on a square metre sub-sample, and grains per ear was calculated from yield, TGW, and ear count data.

Results for Soissons in 2017 are unavailable due to contamination of the plots with seed from another variety. The missing results were entered as missing values as part of the statistical analysis of the data, and the predicted values generated by the ANOVA model are presented throughout (marked *). For any data averaged across years, genotypes, or nitrogen treatments, the predicted means from the ANOVA model are presented, which include corrections made for the missing values for Soissons in 2017.

4.3.1. Yield

Grain yield was measured on grain sampled from the 3m × 9m field trial plots, from which yield in tonnes per hectare (t/ha) was extrapolated. Yield data for all combinations of year, genotype, and nitrogen supply is presented in table 4.5.

Analysis of the yield data from the WGIN diversity field experiment identified a significant two-way interaction between genotype and year of experiment ($F_{5,32}$=11.94, P<0.001, SED=0.364 on 9.82 DF). However, this interaction is primarily describing the high yield of Hereward relative to other genotypes in 2016 (see figure 4.6), and whilst this may prove interesting, it is unlikely to be the best descriptor of the data as a whole. Therefore the individual effects of year, nitrogen, and genotype are also presented.

Yield was significantly increased by the high nitrogen (350kg-N/ha) treatment ($F_{1,6}$=131.53, P<0.001, SED=0.279 on 6 DF) with an average yield increase of 39%, from 8.16 to 11.36t/ha between the 100 and 350kg-N/ha nitrogen treatments respectively (figure 4.7a). Genotype was also found to have a strong effect on yield ($F_{3,32}$=50.82, P<0.001, SED=0.115 on 32DF): Istabraq was the highest yielding with a yield of 10.41kg/a, followed by Hereward with 10.07kg/ha, Cadenza with 9.37kg/ha, and Soissons with 9.18kg/ha (figure 4.7b). Finally, yield was also significantly different between years ($F_{1,6}$=17.12, P=0.003, SED=0.321 on 6DF), with 2015 having the highest average yield at 10.74kg/ha, followed by 2016 with 9.67kg/ha, and 2017 showing a 17% decrease compared to 2015 with 8.87kg/ha (figure 4.7c).

Table 4.5: Full grain yield dataset from the WGIN diversity field trial experiment. Grain yield expressed as t/ha, and is averaged across experimental blocks. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		Yield (t/ha)		
Nitrogen input	Genotype	2015	2016	2017
100kg-N/ha	Cadenza	8.21	7.25	7.46
	Hereward	8.83	8.78	7.78
	Istabraq	9.84	8.26	8.65
	Soissons	8.86	6.75	7.23*
350kg-N/ha	Cadenza	12.49	10.90	9.90
	Hereward	12.28	12.64	10.13
	Istabraq	13.37	11.89	10.46
	Soissons	12.04	10.87	9.36*

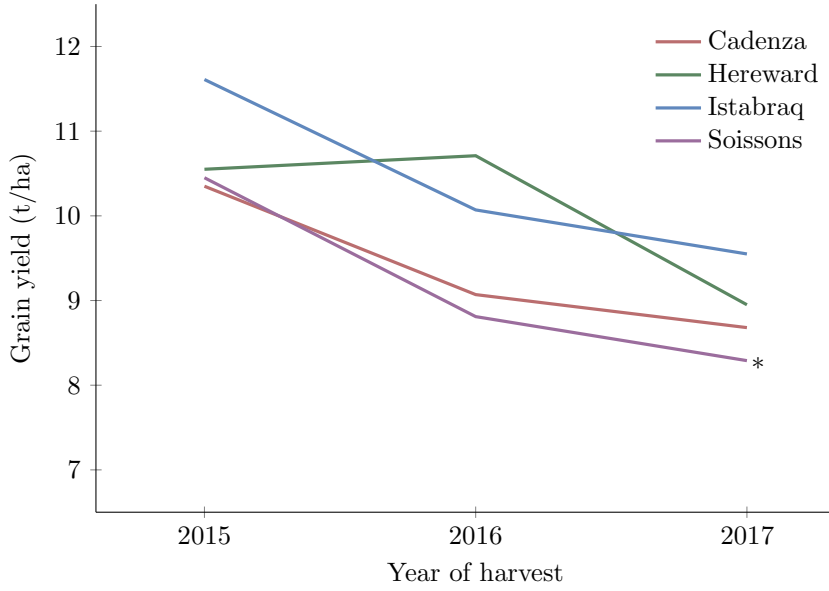
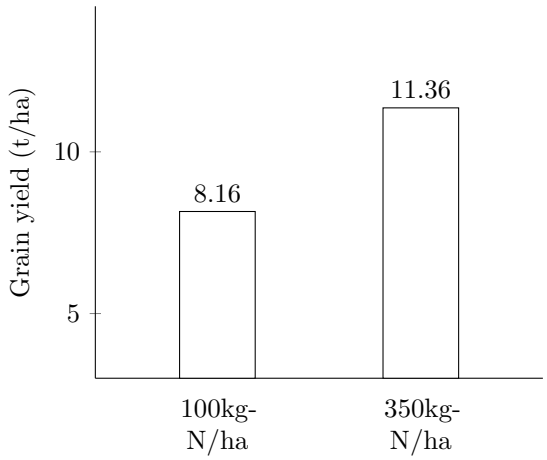
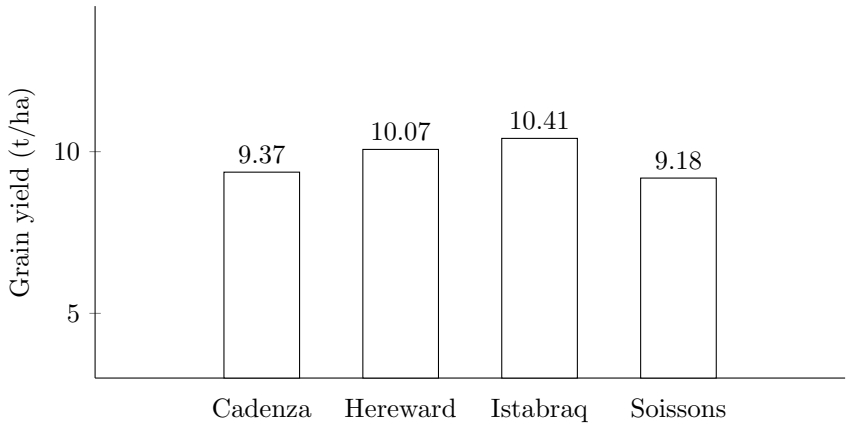


Figure 4.6: **The interaction between genotype, year, and grain yield in the WGIN diversity field experiment.** Grain yield by genotype and year from the WGIN diversity field experiment, averaged across nitrogen treatments. Figure represents the two-way interaction between genotype and year of experiment identified by ANOVA, which largely describes the high yield of Hereward in 2016. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *). LSD (at the 5% level) of 0.41 when comparing means within the same year, and 0.81 for all other comparisons.

(a)



(b)



(c)

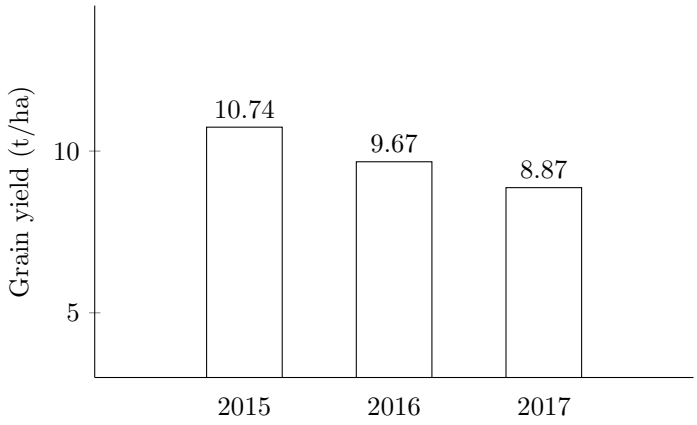


Figure 4.7: **Grain yield is increased by nitrogen fertilisation, varied between genotypes, and was lowest in the hottest year.** Grain yield data from the WGIN diversity field experiment presented by the individual effects of (a) nitrogen treatment, LSD (at the 5% level) of 0.684; (b) genotype, LSD of 0.234; and (c) year, LSD of 0.785

4.3.2. Thousand grain weight

Thousand grain weight was measured on a sub-sample of the grain harvested at maturity, and as with yield, was measured on grain bench-dried to 15% moisture.

Year of harvest, genotype, and nitrogen application rate all interacted to determine TGW ($F_{5,31}$=3.71, P=0.010, SED=1.309 on 27.8 DF) (figure 4.8). This three-way interaction is largely describing the response of Soissons to year of harvest, as well as the different responses to nitrogen input. Over the three years, TGW was generally greatest in 2016, and lowest in 2017. However, for Soissons, TGW was lower in 2016 than in 2015, a response that wasn't observed in any other genotype. Looking at the genotypes individually, Cadenza had the highest average TGW at 42.1g, Hereward and Istabraq had TGWs of 40.4g and 40.6g respectively, whilst Soissons had the lowest TGW at 35.0g. The effect of nitrogen supply was comparatively small, and varied between combinations of genotype and year. Whilst increased nitrogen input generally *decreased* TGW, this response was not consistent, and indeed wasn't universal across all treatment combinations. By comparing the predicted means with the LSD at the 5% level, it is evident that TGW was only significantly decreased under the high nitrogen treatment in Istabraq in 2016, and Hereward in 2017, and that no instances of increased in TGW due to nitrogen treatment were significant.

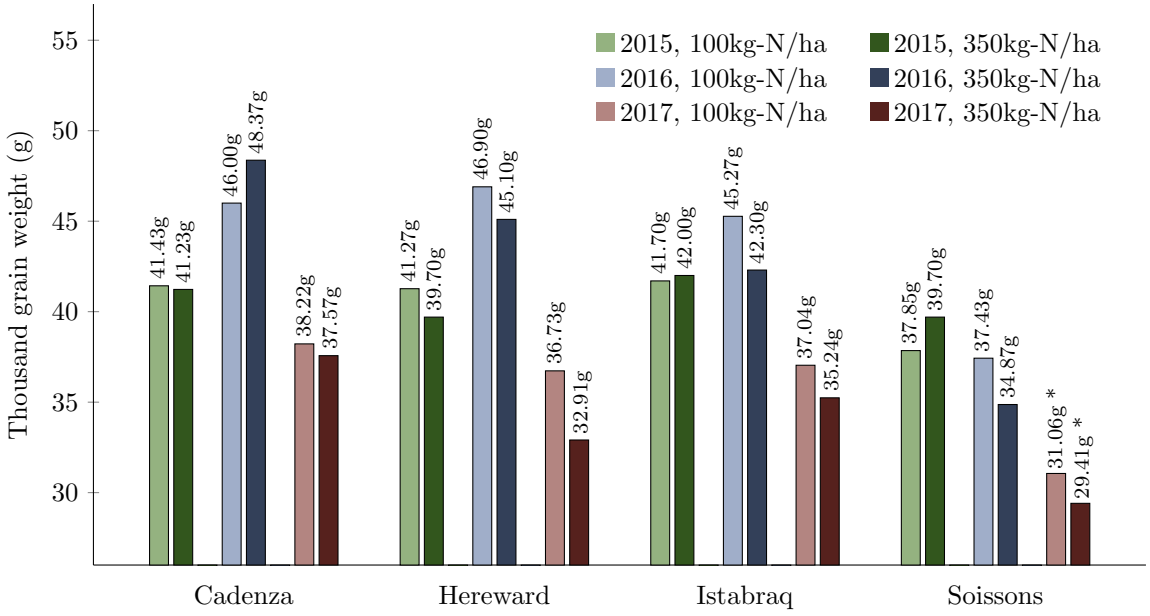


Figure 4.8: **Thousand grain weight is affected by nitrogen input, and varies between genotypes and year.** TGW data from the WGIN field experiment shown as individual treatment combinations of genotype, nitrogen, and year of experiment. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *). LSD (at the 5% level) of 1.94 within the same combination of year and nitrogen treatment, 2.66 within the same year, or combination of year and genotype, and 2.68 for all other comparisons.

.

Table 4.6: Full ear count dataset from the WGIN diversity field trial experiment. Ear count measured on a one square metre sub-sample, and the average value over three experimental blocks is presented. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		Ears per m^2		
Nitrogen input	Genotype	2015	2016	2017
100kg-N/ha	Cadenza	361	417	298
	Hereward	460	559	393
	Istabraq	385	422	333
	Soissons	559	399	386*
350kg-N/ha	Cadenza	483	529	405
	Hereward	535	539	477
	Istabraq	506	512	493
	Soissons	631	772	643*

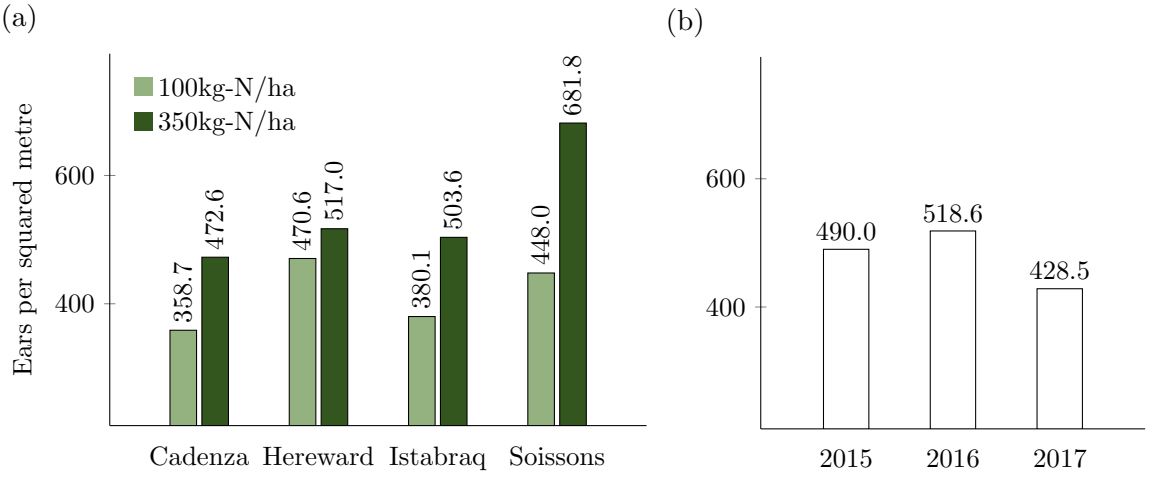


Figure 4.9: **Ear count is increased by nitrogen fertilisation, and was highest in the coolest year.** Ear count data from the WGIN field experiment shown by (a) genotype and nitrogen treatment combination, LSD (at the 5% level) of 78.3 when comparing within the same level of nitrogen, and 71.9 for all other comparisons; and (b) year, LSD of 23.2.

.

4.3.3. Ear count and grain per ear

The number of ears in a one square metre sub-sample was counted, and using grain yield and TGW measurements, an average number of grains per ear was calculated. Data for the number of ears are presented in table 4.6, and for the number of grain per ear in table 4.7.

Both genotype and nitrogen input interacted to determine ear count ($F_{3,31}$=4.09, P=0.015, SED=35.46 on 36.56 DF) (figure 4.9a), with Cadenza, Istabraq, and Soissons all showing a significantly higher ear count under the high (350kg-N/ha) nitrogen treatment. Whilst the ear count for Hereward wasn't significantly increased by the high nitrogen treatment, it could equally be stated that the ear count wasn't *reduced* by limiting nitrogen application, since

Table 4.7: Full dataset for grain per ear calculations from the WGIN diversity field trial experiment. Grain per ear calculated from ear count and TGW data. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		Grain per ear		
Nitrogen input	Genotype	2015	2016	2017
100kg-N/ha	Cadenza	47.1	32.8	55.8
	Hereward	39.7	31.5	46.1
	Istabraq	52.4	38.8	59.7
	Soissons	34.9	38.4	50.5*
350kg-N/ha	Cadenza	53.2	36.9	56.9
	Hereward	49.7	44.4	56.0
	Istabraq	53.5	47.5	51.8
	Soissons	40.9	34.5	45.1*

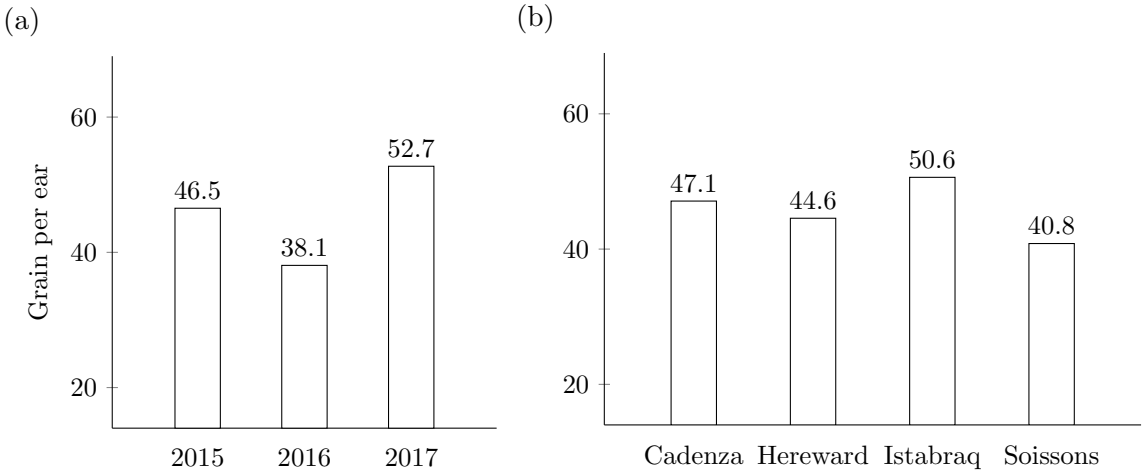


Figure 4.10: **Ears contained more grain in hotter years, and ear size varied between genotypes.** Grain per ear data from the WGIN diversity field experiment grouped by (a) year, LSD (at the 5% level) of 4.9; and (b) genotype, LSD of 5.0.

.

Hereward achieved the highest ear count under the low (100kg-N/ha) nitrogen treatment. Of the four genotypes that saw an increase in ear count due to nitrogen input, Soissons showed the strongest response with a 50% increase in the number of ears, whilst Cadenza and Istabraq showed increases of 32% and 33% respectively. In comparison, the ear count of Hereward varied by only 10% between nitrogen treatments. Ear counts were also significantly different between years ($F_{2,6}$=47.14, P<0.001, SED=9.49 on 6 DF) (figure 4.9b), with the highest ear count recorded in 2016 with 519 ears per metre squared, 490 in 2015, and the lowest count of 429 in 2017.

The number of grain per ear was significantly different between years ($F_{2,6}$=27.44, P<0.001, SED=1.987 on 6 DF) (figure 4.10a), with the most grain per ear recorded in 2017 with 52.7, 46.5 in 2015, and the fewest in 2016 with 38.1 grain per ear. Grain per ear also varied between genotypes ($F_{3,31}$=5.73, P=0.003, SED=2.435 on 31 DF) (figure 4.10b) with mean grain per

ear counts of 50.6 for Istabraq, 47.1 for Cadenza, 44.6 for Hereward, and 40.8 for Soissons. Whilst not significant at the 5% level, there was some evidence for an effect of nitrogen input ($F_{1,6}$=5.46, P=0.058), whereby higher nitrogen input increased grain per ear from an average of 44.0 to 47.5 grain.

4.4. **Discussion**

4.4.1. Controlled-environment experiment

The controlled-environment experiment used liquid nutrient solutions applied during vegetative development (prior to anthesis) in combination with a post-anthesis elevated temperature treatment to identify the combined effect that nitrogen supply and temperature have on the morphology of the mature grain. Both post-anthesis temperature and nitrogen input had significant effects on the grain morphology, each contributing to changes in yield by different mechanisms. Whilst the yield data demonstrates the ultimate effect of these treatments on grain yield, the data for TGW, grain count, and for grain size describe the means by which yield is determined.

SPAD data for leaf chlorophyll content was collected shortly after anthesis to confirm that the two nutrient solutions applied prior to anthesis were different enough to have a small, but statistically significant effect on the nitrogen status of the plants. Both the low and the high nitrogen treatments had average SPAD values above 50, with a difference between the two treatments of 1.64. By comparing the collected SPAD values with values from other studies (Islam *et al.* 2014; Monostori *et al.* 2016), it is clear that the difference between the two nitrogen treatments used in this study was minimal, and also that the plants given the low-nitrogen treatment were by no means deficient in nitrogen. The minimal difference between the two nitrogen treatments is likely due to the composition of the potting mix used, which is discussed further in section 8.2.

Grain yield was greatly reduced by high temperatures during grain filling, and to a lesser extent by limiting the nitrogen supply prior to anthesis, confirming previous reports (Thorne *et al.* 1987; Mitchell *et al.* 1993). Of the two temperature treatments, plants grown under control temperature showed a greater yield response to nitrogen input, with plants grown at high temperature showing a minimal increase in grain yield when nitrogen supply was greater. This differential response to nitrogen input suggests that the reduction in yield observed when temperatures are high during grain-filling cannot be reversed by increasing nitrogen supply to the crop.

Elevated temperature during grain filling resulted in a decrease in TGW, as previously described by Sofield *et al.* (1977). The reduction in TGW under high post-anthesis temperature was of a comparable magnitude to the reduction in yield under the same conditions, and so this yield response can largely be explained by a reduction in grain size. To understand the degree to which TGW was reduced by high temperature, comparisons can be made to the data collected from the WGIN diversity field experiment. The clearest comparison to make

is between the TGW measurements taken under the high temperature treatment, and those recorded from the hottest year of the field experiment (2017). The hottest year of the field experiment also correlated with the lowest TGW measurements, but even the lowest TGW achieved in the field was still approximately 10g higher than the TGW from the elevated temperature treatment in the controlled-environment experiment. This perhaps shows how harsh the high temperature treatment was on growth and development of the plants when compared to typical year-to-year variation. In addition to elevated temperature, lower nitrogen input also reduced TGW, albeit to a lesser extent. Furthermore, when compared with the associated reduction in yield observed under lower nitrogen input, the reduction in TGW is minimal in comparison, which suggests that TGW alone cannot explain the differences in grain yield observed between the two nitrogen treatments used in this experiment.

The analysis of the TGW data showed that the reduction in yield caused by high temperature during grain-filling could largely be explained by a reduction in grain size, whilst the reduction associated with reduced nitrogen input could not. Grain count was not significantly different between the two post-anthesis temperature treatments, confirming the assertion that the differences in yield were predominantly due to changes in grain weight. This result was expected, since whilst high temperature is known to reduce grain set, the mechanisms by which this occurs is through either reduced tiller (and therefore ear) production (Thorne *et al.* 1987), or by inducing sterility when extremely high temperatures are experienced in the days immediately following anthesis (Tashiro *et al.* 1990): our temperature treatment was only applied post-anthesis, and used a relatively mild temperature. Grain count was, however, reduced under the low nitrogen treatment. This result shows that lower levels of nitrogen input reduce yield without reducing the grain weight: the *number* of grain produced by a plant receiving less nitrogen is decreased. Since grain count was derived from yield and TGW data collected at harvest, it is impossible to dissect what caused this reduction in grain count, *i.e.* whether it was a reduction in the amount of ear-bearing tillers, a reduction in the size of each ear, or a combination of both of these factors.

Measurements taken on the dimensions of the mature grain provide further information on how the morphology of the grain is altered by different nitrogen and temperature conditions. Grain area measurements mirror the TGW results, with high temperature during grain-filling resulting in a large decrease in grain area, and limiting nitrogen application resulting in a smaller decrease. However, the effects observed in the grain area data are less pronounced than those seen in the TGW data, which suggests that the changes in TGW are due to a change in both grain size, *and* grain density. Interestingly, the dimensional measurements of length and width showed different responses to the nitrogen and temperature treatments used: grain was longer when provided with more nitrogen during development, and grain were narrower when temperatures during grain-filling were elevated.

4.4.2. WGIN diversity field trial experiment

The aim of the WGIN diversity field trial experiment was to investigate the impact that nitrogen supply and climate have on grain yield and morphology in different genotypes. As well as yield and TGW, the number of ears produced was also recorded. From this information is was possible to calculate the size of each ear, providing more information on the determinants of grain yield.

Initial analysis of the yield data from the WGIN diversity field experiment found a two-way interaction between year of harvest and genotype, which signifies that different genotypes responded differently to different years. However, this interaction was found to be describing the high yield of Hereward in 2016 when compared to other genotypes (see figure 4.6). Whilst the other three genotypes sampled during this experiment showed a decrease in yield between 2015 and 2016, Hereward showed an increase that was observed for both the low (100kg-N/ha) and high (350kg-N/ha) nitrogen treatments. It is therefore clear that the conditions of 2016 were uniquely favourable to Hereward. Although the two-way interaction between genotype and yield identified the response of Hereward to the conditions of 2016, it fails to signify any interesting response in the other genotypes sampled. Therefore the individual effects of genotype, nitrogen, and year were also investigated. When looking at the treatments individually, increased nitrogen supply had the greatest effect on yield, with a 40% increase in yield due to increased nitrogen application. This is not surprising, since unlike in the controlled-environment experiment, the difference between 100kg-N/ha and 350kg-N/ha nitrogen treatments was great enough to cause significant phenotypical differences between plants: plants grown under the low nitrogen treatment were smaller, with fewer tillers, and its leaves were a paler shade of green when compared to plants grown with plentiful nitrogen. Yields were also significantly different between the different genotypes sampled, with the NABIM group four feed wheat Istabraq achieving the highest yields, closely followed by Hereward, and with the lowest yields from Soissons and Cadenza. As a feed wheat, Istabraq is low in protein, and due to grain protein deviation, whereby high protein wheats generally yield lower (Simmonds 1995), could be expected to yield higher than the high-protein bread-making wheats. The high average yield of Hereward can be partly explained by its high performance in 2016, but perhaps also by its tendency to show high grain protein deviation (Monaghan *et al.* 2001), achieving a higher yield than could be expected when considering its protein content. Grain yield also varied between years, with 2015 emerging as the highest yielding year, followed by 2016 and 2017. This is most likely due to the differences in the climate experienced during grain-filling, and is discussed in more detail in section 4.4.2.1.

With regards to TGW, the biggest determining factor was year of experiment. The largest grain were generally recorded in 2016, with the only clear exception to this trend being Soissons, which had larger grain in 2015 than in 2016. However, when viewed alongside the yield data, it is perhaps Cadenza and Istabraq that stand out, since both of these genotypes had larger grain in 2016 compared to 2015, whilst achieving lower yields. In this respect,

Hereward and Soissons are alike, in that the TGW data was more closely correlated to grain yield. Therefore it may be that grain size (as TGW) is a stronger determinant of yield in Hereward and Soissons than in Cadenza and Istabraq. Whilst clear differences were evident between the TGW measurements of different genotypes, and between different years, the effect of nitrogen was comparatively weak. In contrast to the controlled-environment experiment, increased nitrogen input did not result in an increase in grain size. Rather, TGW was generally reduced by the high nitrogen treatment, although only statistically significant at the 5% level for Istabraq in 2016 and Hereward in 2017. Although this is in contrast to the findings of the controlled-environment experiment, it is not unheard of for nitrogen input to have this effect (Kindred *et al.* 2008), with certain genotypes more likely to show a reduction in TGW accompanied by an increase in yield under higher levels of nitrogen input. The inclusion of Cadenza in both the controlled-environment and the WGIN diversity field experiment allows for some direct comparisons to be made between the two experiments. In the controlled-environment experiment, higher nitrogen input resulted in an increase in TGW, an effect that was only observed in year 2016 of the field experiment, and was not statistically significant.

In contrast to TGW, the number of ears counted per square metre was greatly increased by the high nitrogen treatment, with an increase of 50% seen in Soissons. Whilst Soissons stands out due to its large response to nitrogen fertiliser input, Hereward showed a much smaller response, which could perhaps more accurately be described as a minimal drop in ear count under lower nitrogen conditions. As with number of ears, nitrogen application increased the number of grain within each ear, the combined effect of which being more grain being produced per unit of area under the high nitrogen application rate. However, in this instance there was no interaction between genotype and nitrogen application rate, with different genotypes showing a comparable response to nitrogen input. Among the four genotypes, the highest (Istabraq) and lowest (Cadenza) yielding varieties had comparable ear count results. When compared with the grain-per-ear data, it is clear that Cadenza's smaller ears are resulting in a decreased yield. The result is a 12% decrease in the amount of grain produced when compared to Istabraq, and even with the higher TGW achieved by Cadenza, this is enough to create a large difference in final yield. Soissons also stands out, with the highest count of ears out of all of the genotypes, but the lowest number of grain per ear. These two results essentially cancel each other out, resulting in a similar number of grain being produced to Istabraq and Hereward (but more than Cadenza), which when combined with a low TGW results in a low yielding crop. Of the three years, 2015 had the highest grain yield, but 2016 had the highest TGW. Again, comparing the number of ears, and the size of these ears, shows that in 2015 more grain were produced over fewer ears, resulting in an increased yield. Again, the effect of year-to-year variation is discussed further in section 4.4.2.1

4.4.2.1. *Effect of climate*

Of the three years sampled as part of the WGIN diversity field experiment, the highest yields were observed in 2015, followed by 2016, and with the lowest yields recorded in 2017. With regards to accumulated thermal time, 2015 and 2016 were comparable, and both accumulated less thermal time than 2017. By only considering accumulated thermal time, it could be predicted that 2016 would have achieved the highest yields, since this was the year with the mildest temperatures during grain-filling. However, the difference in grain yield between 2015 and 2016 is considerable, with the yields in 2016 closer to those obtained in 2017, a much hotter, drier year. The most obvious difference between the climate experienced during grain-filling in 2015 and 2016 is in the amount of accumulated sunlight. In 2016 the amount of accumulated sunlight was approximately two-thirds of the sunlight accumulated in 2015, and so it is reasonable to attribute lack of incident radiation to the decrease in yields observed in 2016. With regards to the lowest yields, recorded in 2017, these are doubtlessly a product of the high temperatures and lack of rain during grain-filling.

With regards to TGW, the highest values were recorded in 2016, closely followed by 2015, and with the lowest recorded in 2017. Of the meteorological information recorded, accumulated thermal time may best explain the TGW results over the three years. In 2016, the amount of accumulated thermal time was the lowest, which would allow for slower grain development, resulting in a larger grain, an effect clearly evident in the results of the controlled-environment experiment. An exception to this trend, however, is observed in the TGW measurements for Soissons. Soissons was the only genotype for which the TGW measurements from 2015 and 2016 don't correlate with the total accumulated thermal time during grain-filling. It could therefore be possible that Soissons was particularly affected by the lack of sunlight during 2016, and that this resulted in a decrease in the size of the grain produced.

Chapter 5: Grain protein distribution results

5.1. Introduction

The endosperm is the central storage organ of the wheat grain, rich in starch and protein, and is the tissue from which white flour is produced. Within the endosperm, there are clear gradients in the distribution of protein, and these gradients are known to be both quantitative and qualitative (Tosi *et al.* 2011; He *et al.* 2013; Wan *et al.* 2014), *i.e.* different proteins show different distribution gradients. During the production of white flour, the wheat grain is milled using a series of rollers and sieves. This milling process breaks apart the endosperm, separating it from the aleurone and bran layers. As a result multiple mill streams are produced, each with different protein content, composition, and baking characteristics (Wang *et al.* 2007; Zhou *et al.* 2018). These individual mill streams are then blended by millers to produce flour with the desired qualities. Therefore identifying the factors that affect the grain protein distribution, and hence the quality of each mill stream, is of great importance. Furthermore, during milling the separation of endosperm from aleurone and bran layers is not complete, and a proportion of the endosperm remains adhered to the removed aleurone layer. It has long been established that protein is concentrated towards the outer layers of the endosperm (Cobb 1905; Kent 1966; Ugalde *et al.* 1990b), and as a result a disproportional amount of protein is removed during the production of white flour. Therefore it would also be beneficial to understand any factors that might impact on the gradient of protein accumulated within the endosperm, since they would directly affect the amount of protein that is lost during milling.

Whilst the grain protein distribution gradient in wheat has been studied for over a hundred years, to date the quantification of these gradients has been a low-throughput process relying on micro-dissection (Cobb 1905; Ugalde *et al.* 1990a; Ugalde *et al.* 1990b), the sub-sampling of microscopy images (Tosi *et al.* 2011), or experimental milling (He *et al.* 2013). The image analysis technique used in this study (described in section 2.5) is the first published high-throughput method for the quantification of wheat grain protein distribution gradients (Savill *et al.* 2018), and uses a custom Python toolbox run in ArcGIS to spatially analyse protein in light-microscopy images. This image analysis technique uses semi-automated maximum-likelihood classification to detect and measure protein within an image, allowing for the quantification of protein concentration across the endosperm. Furthermore, individual measurements are taken for the area and position (relative to the aleurone layer) of each protein body in order to describe the size-distribution of protein bodies within the endosperm.

Previous studies in wheat (He *et al.* 2013; Wan *et al.* 2014) have shown a link between grain protein distribution and nitrogen supply, with increased nitrogen input resulting in an increase in the gradient in protein between outer and inner endosperm. Likewise, the distribution of grain protein is often different between different genotypes (He *et al.* 2013). Whilst work in

barley has shown that day-length can influence the distribution of Hordein storage proteins (Holopainen *et al.* 2012), an environmental effect on the protein distribution in wheat is yet to be identified. With regards to protein body size-distribution, little work has been completed to date, however, drought has been shown to increase the difference in average protein body size between outer and inner endosperm tissue (Chen *et al.* 2016).

To investigate the effects of temperature during grain-filling, nitrogen fertiliser regime, genotype, and year-to-year climate variations in the field, two experiments were completed. The first was a controlled-environment experiment, where British spring bread-making wheat Cadenza was subject to either a high (28°C) or control (20°C) temperature treatment during grain-filling, and was supplier with either high or low nitrogen nutrient solution applied periodically during vegetative growth. The second experiment was based on samples taken from the WGIN diversity field trial experiment, with grain collected from four commercial wheat varieties grown under either 100kg-N/ha or 350kg-N/ha over two years (2015 and 2017). Grain was harvested at early and mid to late grain-filling, and light-microscopy images analysed to quantify both the gradient in total protein across the grain, and the size-distribution of individual protein bodies relative to the aleurone layer.

5.2. **Controlled-environment experiment**

In the controlled-environment experiment, wheat was grown to anthesis with either high or low (one-tenth strength) nitrogen input, followed by a control (20°C) or elevated (28°C) temperature treatment applied for the duration of grain-filling. The aim of this experiment was to investigate the effects that nitrogen supply and increased temperature have on the distribution of protein within the wheat grain, and to identify any interactions between these two factors.

Grain from the controlled-environment experiment was sampled at two timepoints during early and mid to late grain-filling to analyse both the protein distribution, and the size-distribution of protein bodies within the endosperm. Sampling timepoints were calculated based on accumulated thermal time (see table 2.2), in an effort to sample grain at a comparable developmental stage regardless of the temperature treatment.

5.2.1. Protein concentration gradients

In the REML analysis of the protein concentration gradient data, a significant four-way interaction was found between post-anthesis temperature treatment, nitrogen supply, sampling timepoint, and the mean distance of each measurement from the aleurone layer (*i.e.* the mid-point of each zone) ($F=3.92$, $P=0.049$). Under all treatment combinations, protein concentration was greatest closest to the aleurone layer, and decreased linearly towards the centre of the grain. The gradient in protein concentration was greater in grain subjected to elevated temperature post-anthesis, with this effect unchanged across sampling timepoints under low nitrogen input, but increasing over time under high nitrogen input. The effect of nitrogen input was smaller than the effect of temperature, with a moderate increase in pro-

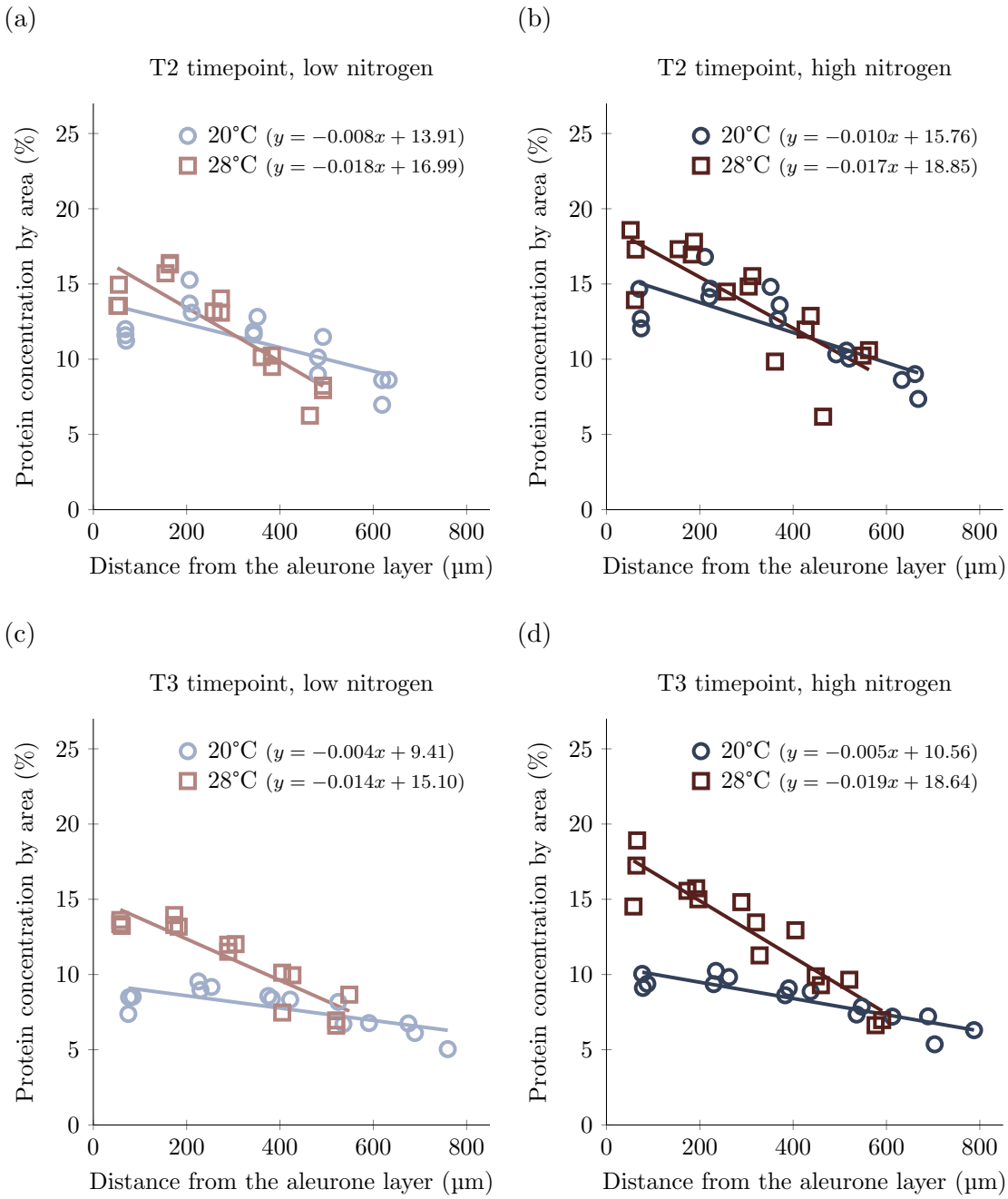


Figure 5.1: **Protein concentration gradients in the wheat endosperm are increased by elevated temperature and high nitrogen input, with the effects increasing over time**. Results of the protein concentration gradient analysis from the controlled-environment experiment at the early (T2) sampling timepoint with (a) low nitrogen input, and (b) high nitrogen input; and from the later (T3) timepoint with (c) low nitrogen input, and (d) high nitrogen input. Temperature treatments are shown within each sub-figure, with open circles and squares representing the control (20°C) and high temperature (28°C) treatments respectively. Trend-lines represent the predictions from the REML analysis, and data-points show the mean protein concentration in each of the five endosperm zones from the three biological replicates.

.

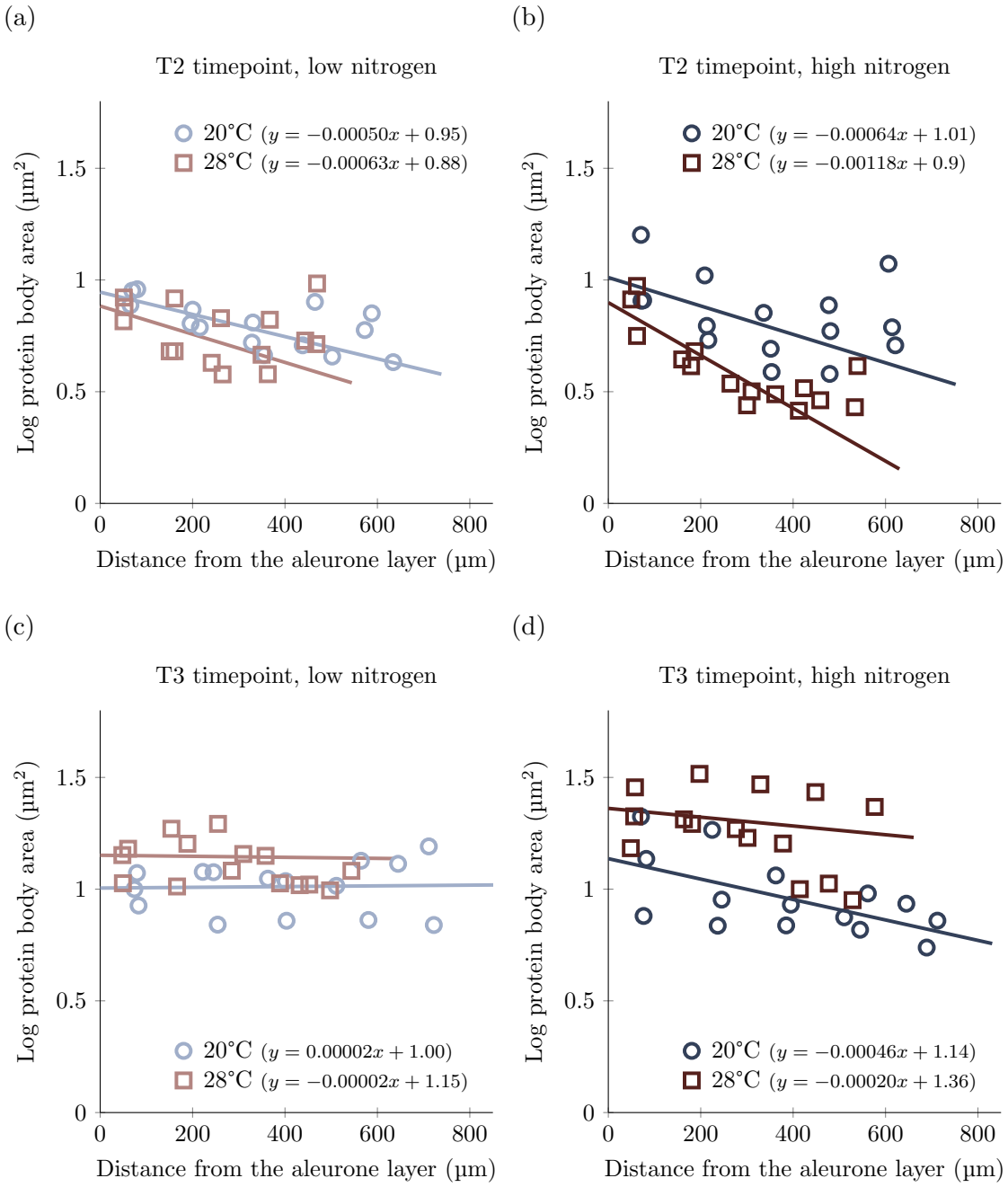


Figure 5.2: **Mean protein body size decreases towards the centre of the grain, with differential effects of nitrogen and temperature over time**. Results of the protein body size-distribution analysis from the controlled-environment experiment at the early sampling timepoint with (a) low nitrogen input, and (b) high nitrogen input; and from the later timepoint with (c) low nitrogen input, and (d) high nitrogen input. Temperature treatments are shown within each sub-figure, with open circles and squares representing the control (20°C) and high temperature (28°C) treatments respectively. Trend-lines represent the predictions from the REML analysis, and data-points show the mean protein body size in each of the five endosperm zones (as used in the protein concentration gradient analysis) from the three biological replicates. Analysis was conducted on log-transformed data, and is presented as such.

.

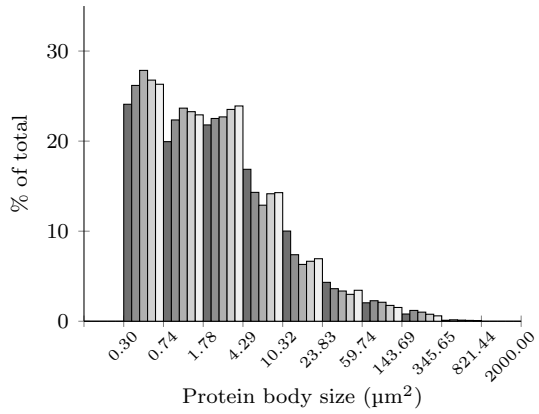
tein concentration gradient in all treatments with the exception of the elevated-temperature treatment sampled at the early timepoint. These results are presented in figure 5.1.

5.2.2. Protein body size-distribution

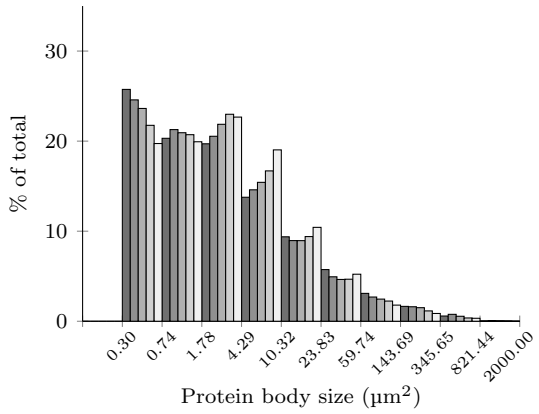
In the protein body size-distribution analysis, a significant four-way interaction was found between post-anthesis temperature, nitrogen supply, sampling timepoint, and distance of protein body from the aleurone layer ($F=213.64$, $P<0.001$). A decreasing gradient in protein body size between the aleurone layer and the central endosperm tissue was observed in all treatment combinations except the control temperature-low nitrogen treatment sampled at the later timepoint, in which protein bodies saw a marginal increase in average size towards the centre of the endosperm. Protein bodies generally increased in size between the early and mid to late grain-filling sampling timepoints, were slightly larger when nitrogen supply was increased, and were affected by elevated temperature differently at each of the two timepoints: at the early timepoint, protein bodies were larger under the control temperature treatment, and at the later timepoint, protein bodies were larger under the elevated temperature treatment. The gradients in protein body size-distribution were also different between the two sampling timepoints. In grain sampled at the early timepoint, the gradient in protein body size-distribution was increased by both elevated temperature and high nitrogen input, with the effect of temperature greater under high nitrogen input, and the effect of nitrogen greater under elevated temperature. However, at the later sampling timepoint the gradients in protein body size-distribution were reduced, with negligible gradients detected in either of the temperature treatments under low nitrogen input that were only slightly increased under high nitrogen input. Under high nitrogen input, the effect of temperature was reversed when compared to grain sampled earlier, with a greater gradient in protein body size-distribution under the *control* temperature regime. The results of the REML analysis of log-transformed data are presented in figure 5.2.

The histograms in figure 5.2 show the breakdown of protein body size across the endosperm for each combination of temperature and nitrogen treatment at each sampling timepoint. The abundance of protein bodies by size are shown for each of the five zones used in the protein concentration gradient analysis (described in section 2.5.1). Although these zones were not used in the protein body size-distribution analysis (since individual measurements of distance from the aleurone layer were recorded for each protein body), they are used here as a convenient means to summarise how protein body sizes vary aross the endosperm. Whilst it is difficult to separate the individual effects of temperature and nitrogen on the protein body size-distribution over the endosperm using these histograms, a few trends emerge which describe both the dataset as a whole and the effects under investigation. These trends are discussed in the discussion towards the end of this chapter (section 5.4.1).

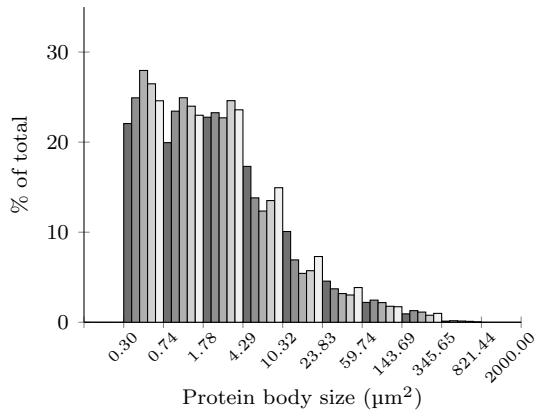
(a)


20°C, low nitrogen, T'2 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

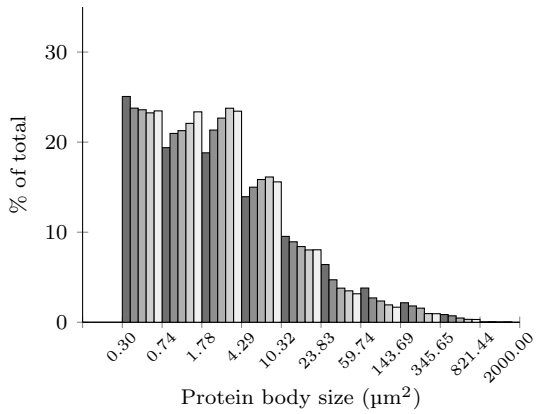
(b)


20°C, low nitrogen, T3 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

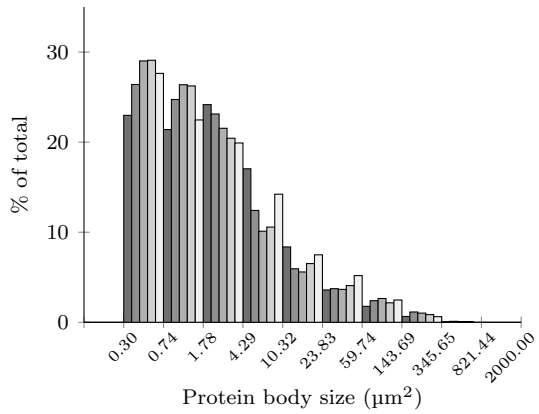
(c)


20°C, high nitrogen, T2 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

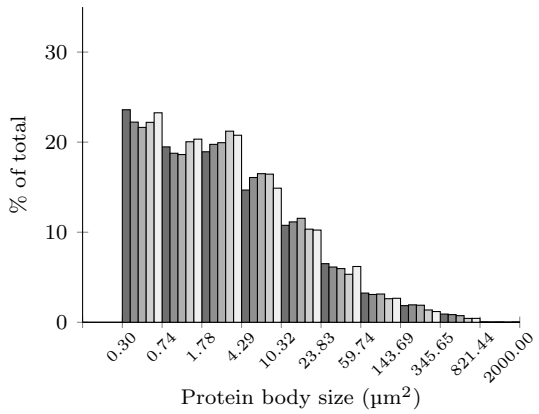
(d)


20°C, high nitrogen, T3 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

(e)


28°C, low nitrogen, T2 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

(f)


28°C, low nitrogen, T3 timepoint
% of total
0
10
20
30
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

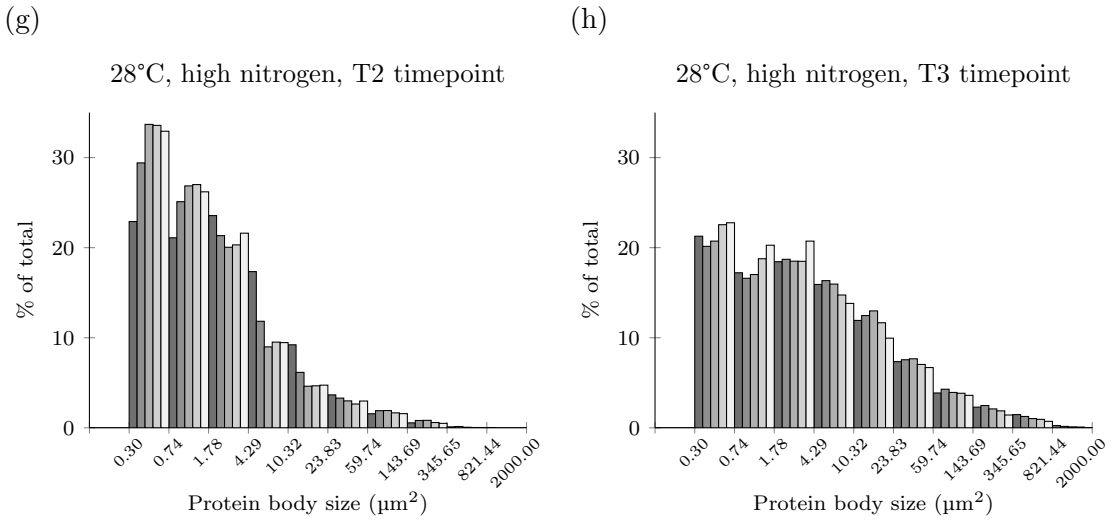


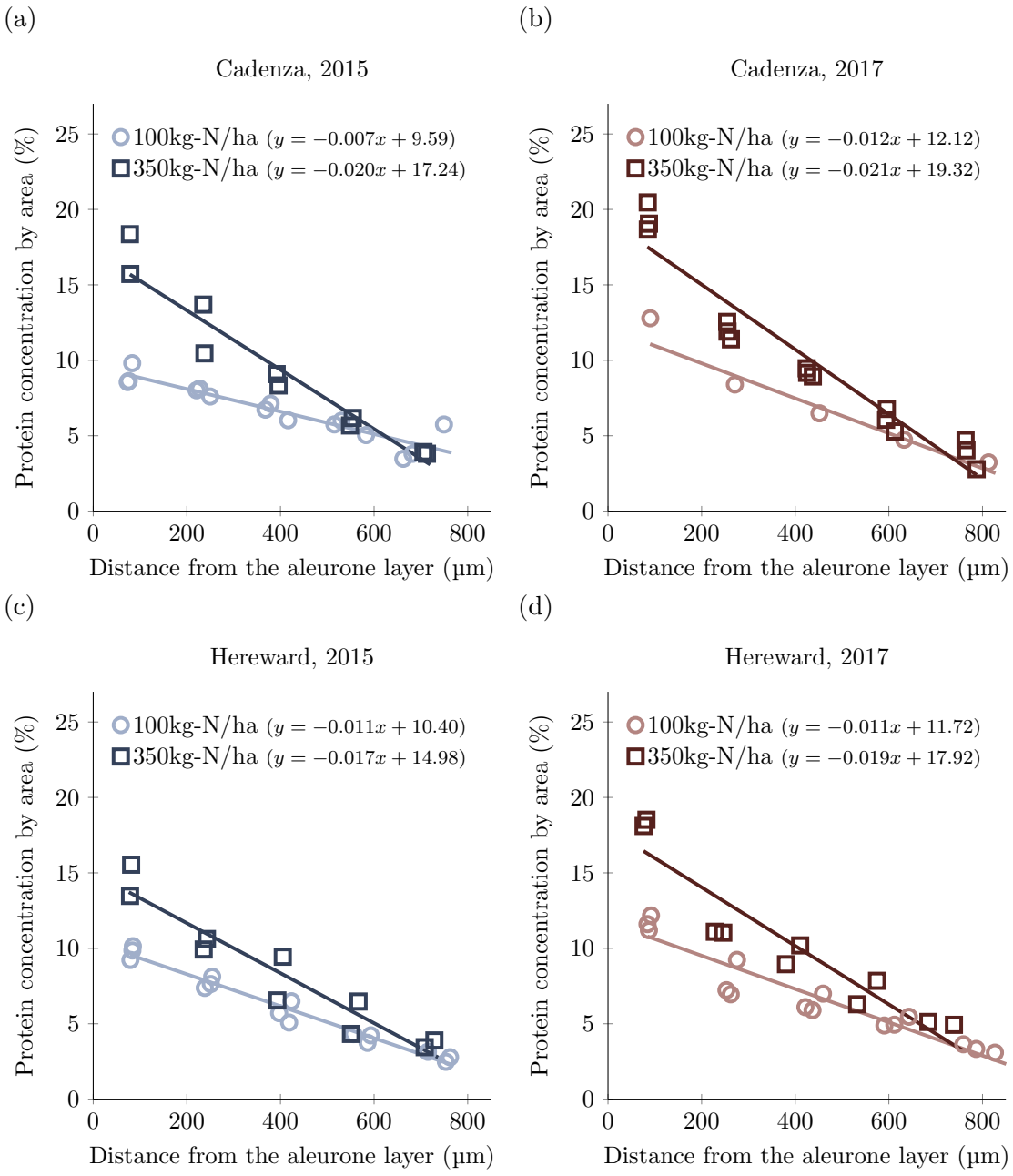
Figure 5.2: **Abundance of different sizes of protein bodies varies across the endosperm, with differential effects of temperature and nitrogen over time**. Histograms showing the frequency of protein body size across the five endosperm zones, from outer in inner endosperm, represented by dark to light grey bars respectively. Each individual sub-figure shows the frequency of protein body sizes in each of the five zones, for a single combination of temperature treatment, nitrogen input, and sampling timepoint: 20°C with low nitrogen input at (a) the early timepoint, and (b) later timepoint; 20°C with high nitrogen input at (c) the early timepoint, and (d) later timepoint; 28°C with low nitrogen input at (e) the early timepoint, and (f) later timepoint; and 28°C with high nitrogen input at (g) the early timepoint, and (h) later timepoint.

.

5.3. **WGIN diversity field experiment**

As part of the WGIN diversity field experiment, grain was collected from four commercial wheat genotypes grown under either 100kg-N/ha or 350kg-N/ha of applied nitrogen fertiliser. The wheat varieties sampled were chosen to represent a variety of genotypes from low-protein feed wheat to high-protein bread-making wheat. Whilst samples were collected in 2015, 2016, and 2017, the samples from 2016 were destroyed during processing. As such, grain from the two years, 2015 and 2017, were analysed to identify the effect that nitrogen input and year-to-year climatic variation have on the grain protein distribution in different wheat genotypes. The aim of this experiment was to build on the findings of the controlled-environment experiment, to demonstrate the effect of nitrogen on the gradients in protein concentration and protein body size-distribution in field-grown wheat, and also to identify any unique responses shown by different genotypes which could be worthy of further investigation.

Unlike in the controlled-environment experiment, only a single sampling timepoint was analysed for material collected from the WGIN field experiment. Grain was sampled at 28 days post anthesis, which represents accumulated thermal time of 8407 GDH in 2015 and 9580 GDH in 2017, compared to the 7224 GDH accumulated by the later sampling timepoint used


(a)
Cadenza, 2015
100kg-N/ha ($y = -0.007x + 9.59$)
350kg-N/ha ($y = -0.020x + 17.24$)
Protein concentration by area (%)
Distance from the aleurone layer (µm)
(b)
Cadenza, 2017
100kg-N/ha ($y = -0.012x + 12.12$)
350kg-N/ha ($y = -0.021x + 19.32$)
Protein concentration by area (%)
Distance from the aleurone layer (µm)
(c)
Hereward, 2015
100kg-N/ha ($y = -0.011x + 10.40$)
350kg-N/ha ($y = -0.017x + 14.98$)
Protein concentration by area (%)
Distance from the aleurone layer (µm)
(d)
Hereward, 2017
100kg-N/ha ($y = -0.011x + 11.72$)
350kg-N/ha ($y = -0.019x + 17.92$)
Protein concentration by area (%)
Distance from the aleurone layer (µm)

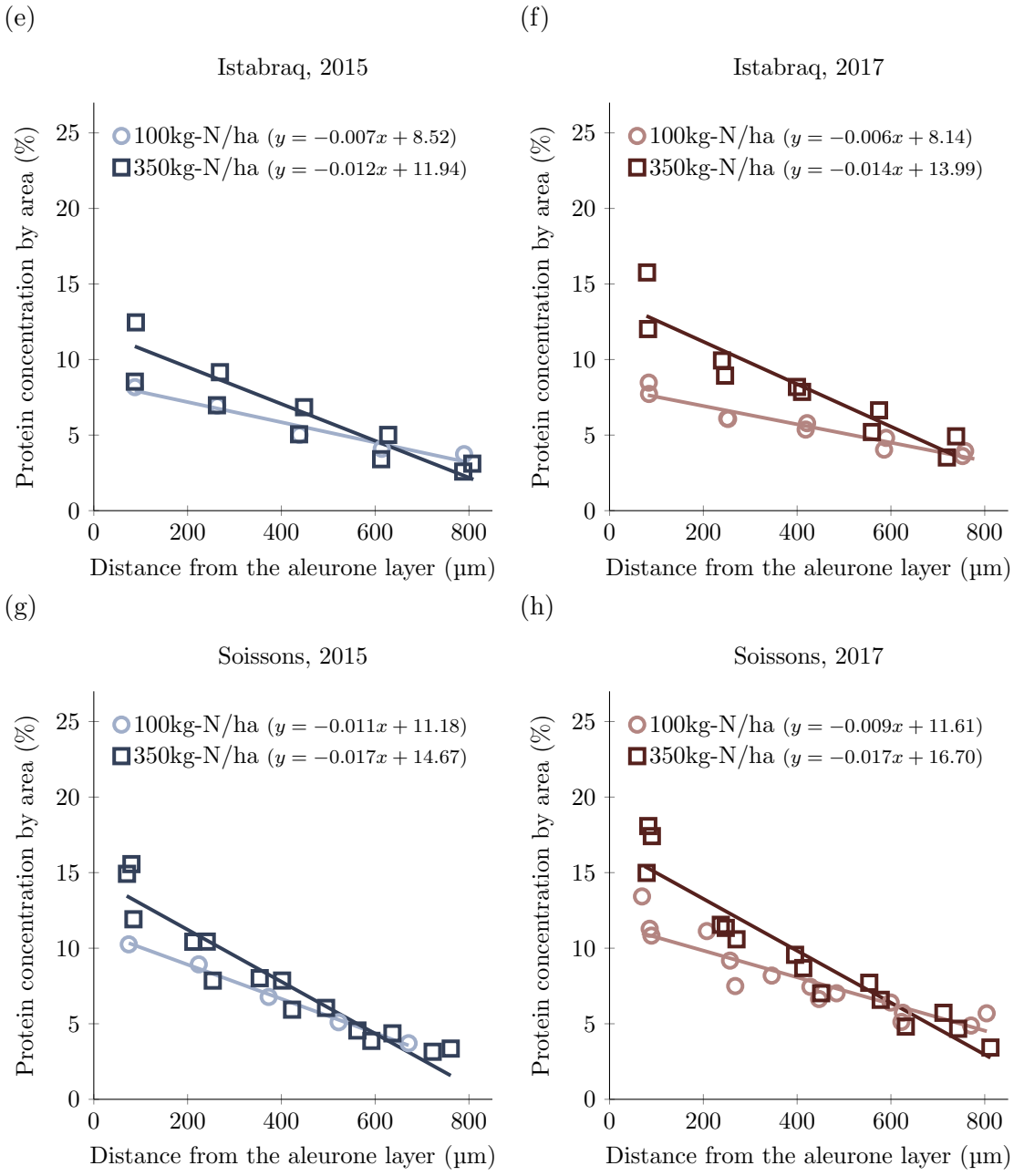


Figure 5.3: **Protein concentration gradients are increased by high nitrogen input in the field, with differential responses from different genotypes over different years**. Results of the protein concentration gradient analysis from the WGIN diversity experiment showing how protein distribution is affected by nitrogen input in four commercial wheat varieties over two years of field experiments. The effect of nitrogen is presented in each sub-figure for Cadenza in (a) 2015, and (b) 2017; Hereward in (c) 2015, and (d) 2017; Istabraq in (e) 2015, and (f) 2017; and Soissons in (g) 2015, and (h) 2017. Nitrogen treatment is represented within each sub-figure, with open circles and squares representing 100kg-N/ha and 350kg-N/ha respectively. Trend-lines represent the predictions from the REML analysis, and data-points show the mean protein concentration in each of the five endosperm zones from each biological replicate (which ranges from one to three).

.

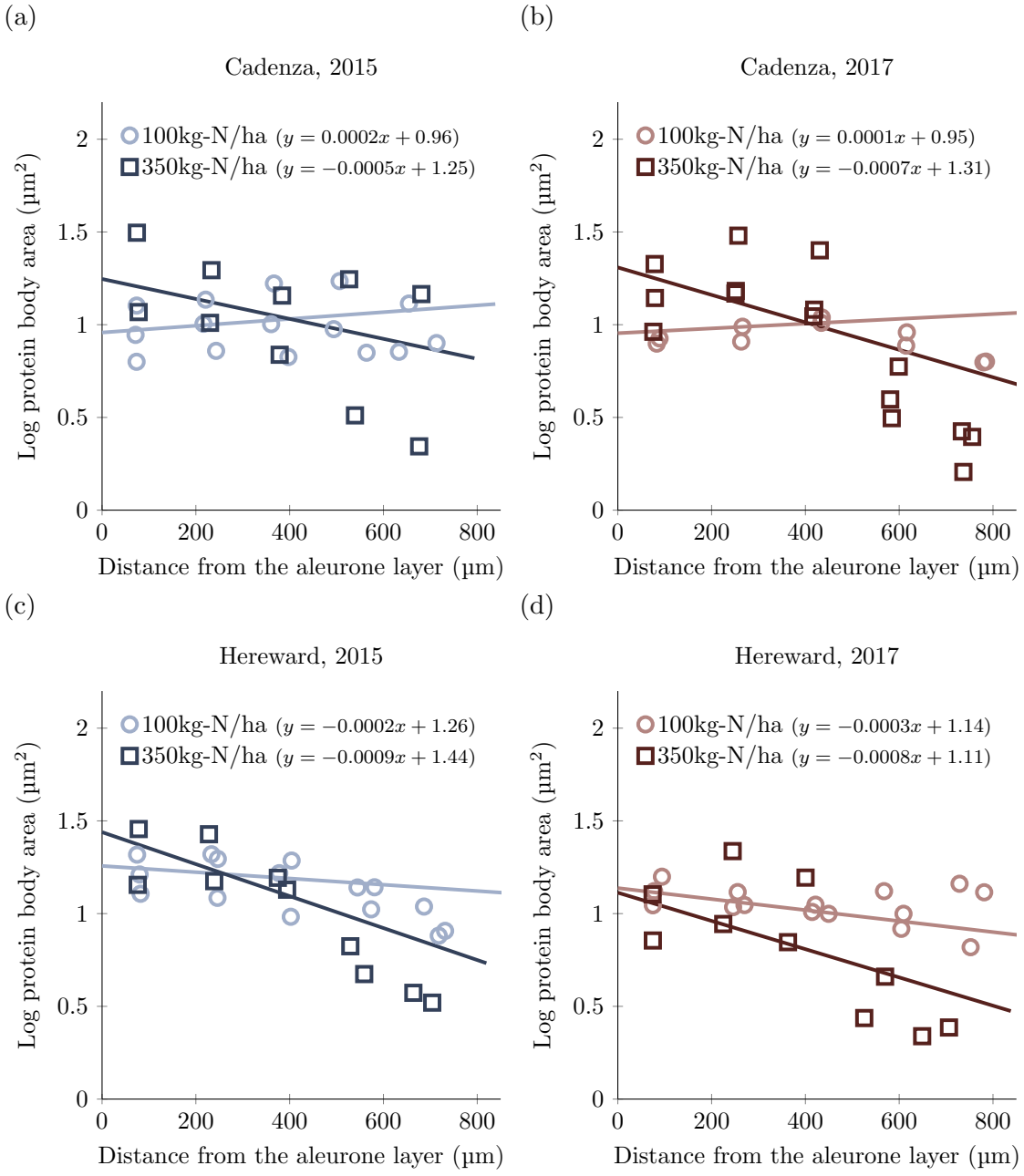
in the analysis of grain from the controlled-environment experiment. Therefore when making comparisons between the two experiments it should be considered that grain sampled from the WGIN field experiment will be more developmentally advanced than the grain analysed as part of the controlled-environment experiment. Likewise, the difference in accumulated thermal time between 2015 and 2017 is considerable, so again it must be kept in mind that the results from each of these years represent grain at different stages of development.

5.3.1. Protein concentration gradients

In the protein concentration gradient analysis, a significant four-way interaction was found between year of experiment, genotype, nitrogen treatment, and the mean distance of each measurement from the aleurone layer ($F=3.71$, $P=0.011$). As in the controlled environment experiment, protein concentration was greatest closest to the aleurone layer, decreased linearly towards the central endosperm tissue, and high nitrogen input was associated with an increase in the gradient of protein across the endosperm. The effect of nitrogen was slightly greater in 2017 compared to 2015 in all genotypes except Cadenza. Of the four genotypes, the steepest gradients were observed in Cadenza, then Hereward, Soissons, and the shallowest gradients in feed-wheat Istabraq. With regards to the response to nitrogen input, Cadenza showed the greatest response, whilst Hereward, Istabraq, and Soissons all showed a similar, weaker response. Differences between years for each genotype were slight, but with Cadenza and Soissons showing the greatest differences between years under low nitrogen input, and Hereward and Istabraq showing the greatest difference under high nitrogen input. The results from this analysis are presented in figure 5.3.

5.3.2. Protein body size-distribution

In the protein body size-distribution analysis, a significant four-way interaction was found between year of experiment, genotype, nitrogen input level, and distance of protein body from the aleurone layer ($F=36.41$, $P<0.001$). In contrast to the results from the controlled-environment experiment there is a lack of a general trend for protein bodies to decrease in size towards the centre of the grain. Rather, in three instances there is an upwards gradient in protein body size, with the average size of protein bodies increasing towards the central endosperm. This upwards gradient is only observed under the low (100kg-N/ha) nitrogen treatment, and was recorded in Cadenza in 2015 and 2017, and Soissons in 2015. In Hereward over both years, and Soissons in 2017, there was a downwards gradient in the size-distribution of protein bodies moving away from the aleurone layer which was increased by high (350kg-N/haN) nitrogen input. Istabraq stands out in this analysis as showing a minimal response to nitrogen with regards to the gradient in protein body size, although it does show a consistent response for larger protein bodies under higher levels of nitrogen input. Between years the response to nitrogen input is relatively consistent between genotypes, with the exception of Soissons. Soissons is unique in that a minimal response to nitrogen was observed in 2015 (both in terms of gradient and average protein body size), whilst in 2017 there is a clear difference between the two nitrogen treatments. The results of this analysis is presented in


(a)
Cadenza, 2015
100kg-N/ha ($y = 0.0002x + 0.96$)
350kg-N/ha ($y = -0.0005x + 1.25$)
Log protein body area (µm²)
Distance from the aleurone layer (µm)
(b)
Cadenza, 2017
100kg-N/ha ($y = 0.0001x + 0.95$)
350kg-N/ha ($y = -0.0007x + 1.31$)
Log protein body area (µm²)
Distance from the aleurone layer (µm)
(c)
Hereward, 2015
100kg-N/ha ($y = -0.0002x + 1.26$)
350kg-N/ha ($y = -0.0009x + 1.44$)
Log protein body area (µm²)
Distance from the aleurone layer (µm)
(d)
Hereward, 2017
100kg-N/ha ($y = -0.0003x + 1.14$)
350kg-N/ha ($y = -0.0008x + 1.11$)
Log protein body area (µm²)
Distance from the aleurone layer (µm)

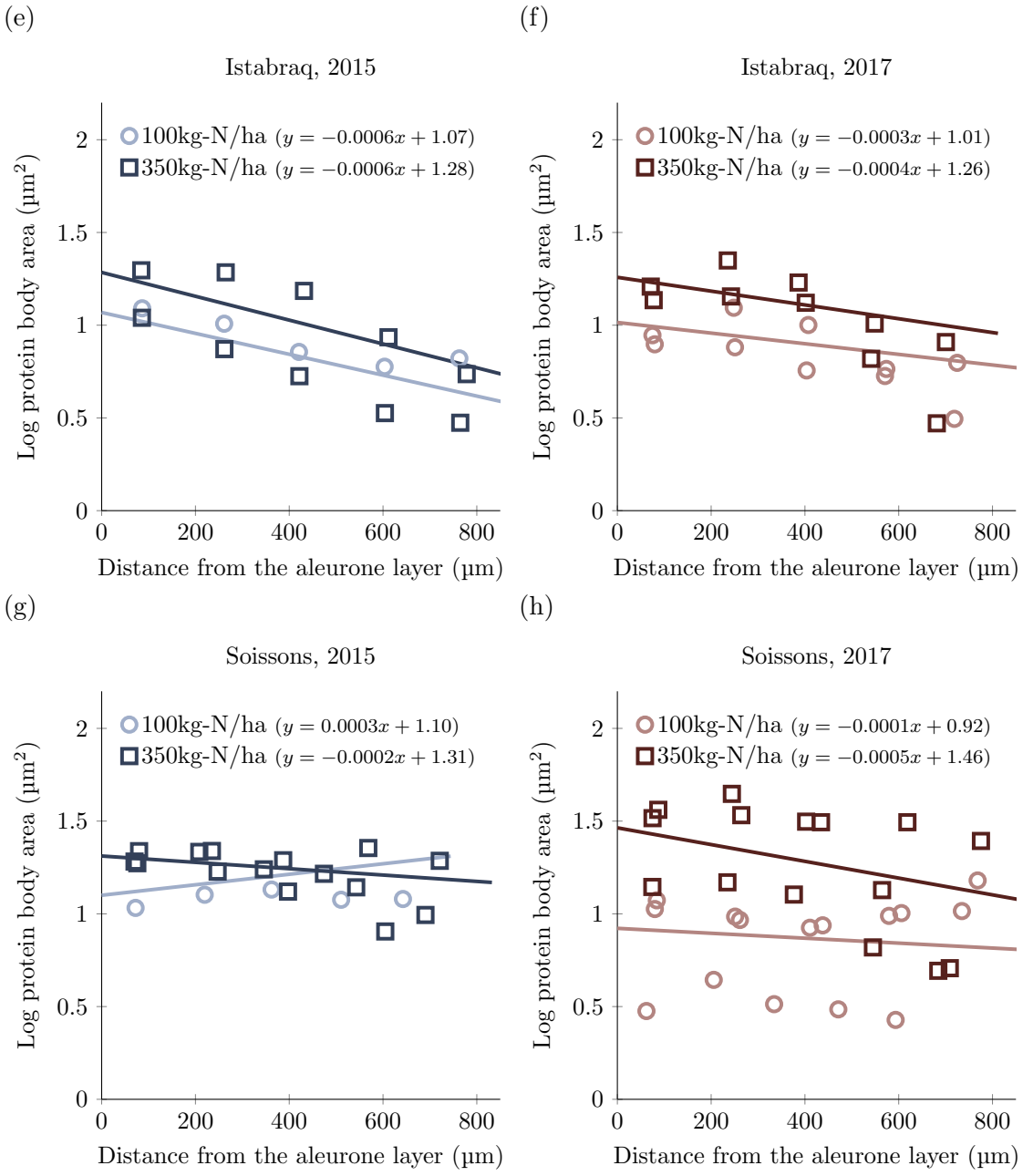
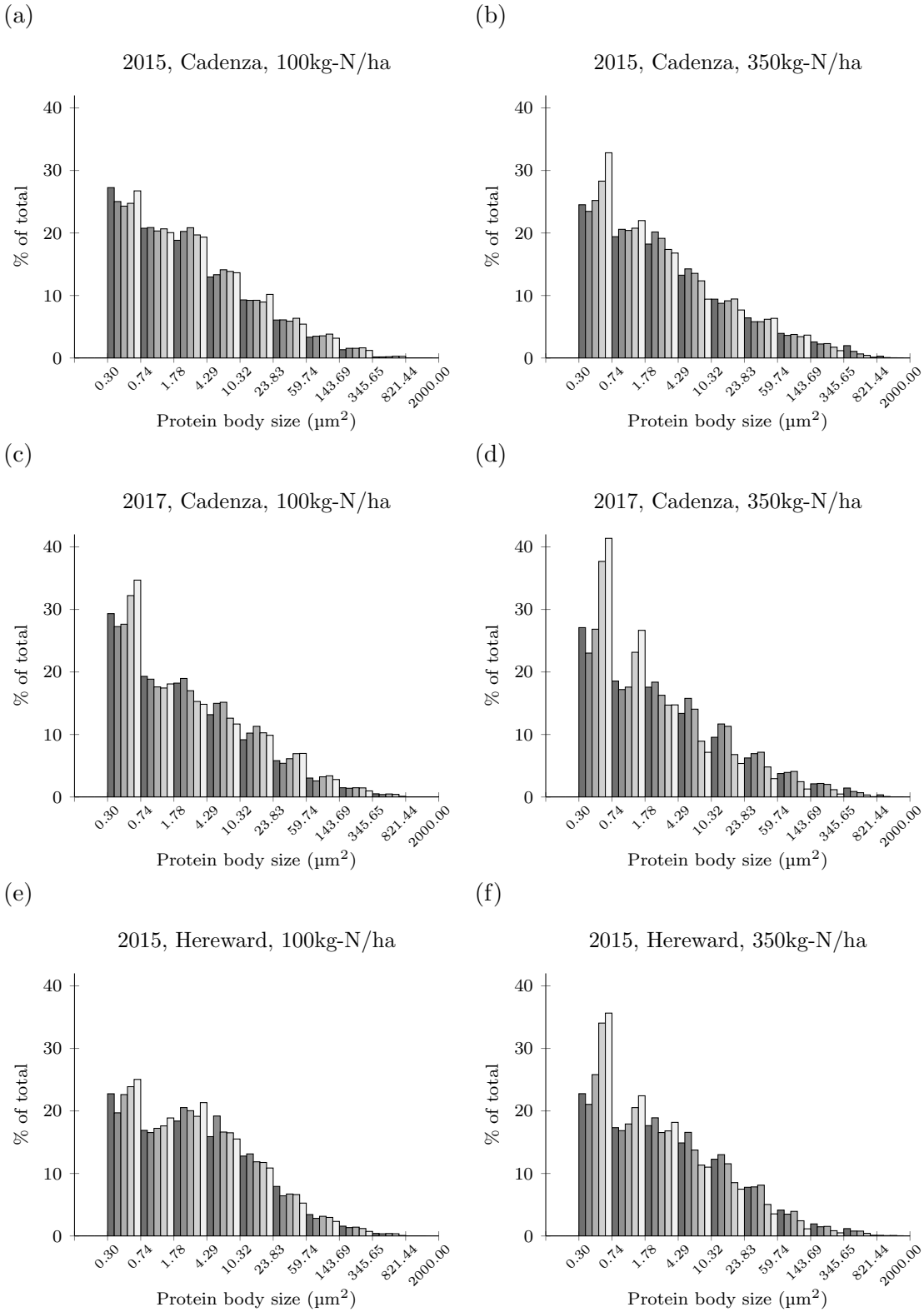
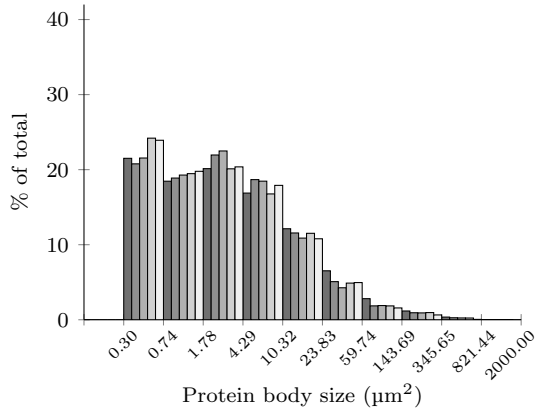


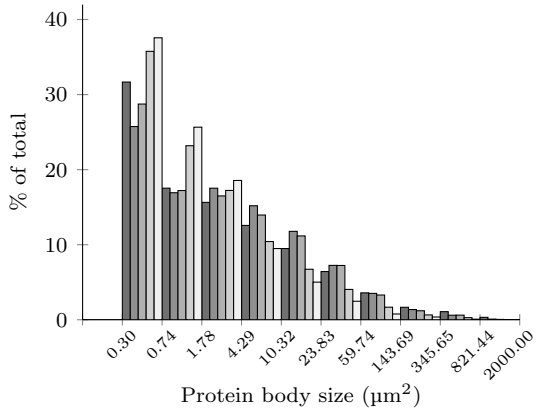
Figure 5.3: **Gradients in protein body size-distribution are determined by genotype, with differential responses to nitrogen input and year of experiment**. Results of the protein body size-distribution analysis from the WGIN diversity experiment showing how protein body size-distribution is affected by nitrgeon input in four commercial wheat varieties over two years of field experiments. The effect of nitrogen is presented in each sub-figure for Cadenza in (a) 2015, and (b) 2017; Hereward in (c) 2015, and (d) 2017; Istabraq in (e) 2015, and (f) 2017; and Soissons in (g) 2015, and (h) 2017. Nitrogen treatment is represented within each sub-figure, with open circles and squares representing 100kg-N/ha and 350kg-N/ha respectively. Trend-lines represent the predictions from the REML analysis, and data-points show the mean protein body size in each of the five endosperm zones (as used in the protein concentration gradient analysis) from each biological replicate (which ranged from one to three). Analysis was conducted on log-transformed data, and is presented as such.


(a)
2015, Cadenza, 100kg-N/ha
(b)
2015, Cadenza, 350kg-N/ha
(c)
2017, Cadenza, 100kg-N/ha
(d)
2017, Cadenza, 350kg-N/ha
(e)
2015, Hereward, 100kg-N/ha
(f)
2015, Hereward, 350kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

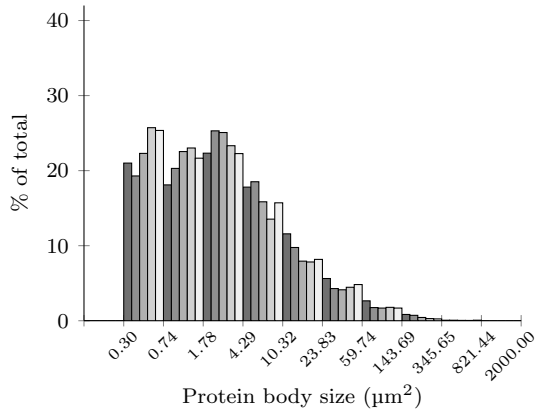
(g)


2017, Hereward, 100kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

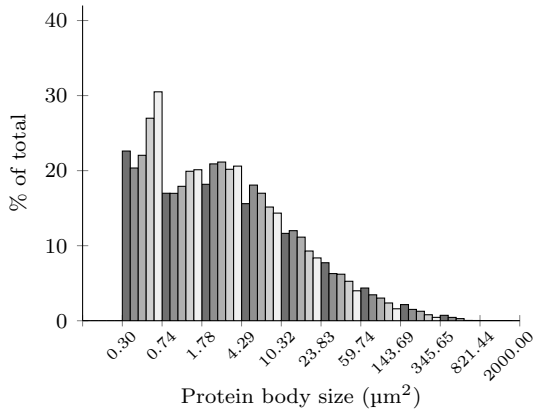
(h)


2017, Hereward, 350kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

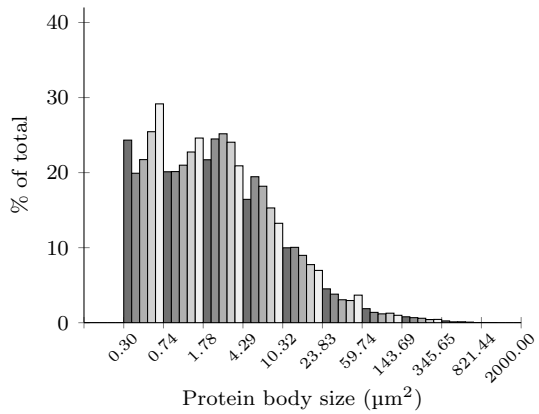
(i)


2015, Istabraq, 100kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

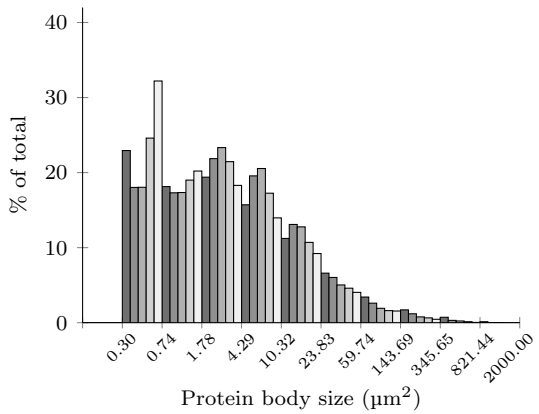
(j)


2015, Istabraq, 350kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

(k)


2017, Istabraq, 100kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

(l)


2017, Istabraq, 350kg-N/ha
% of total
0
10
20
30
40
0.30
0.74
1.78
4.29
10.32
23.83
59.74
143.69
345.65
821.44
2000.00
Protein body size (µm²)

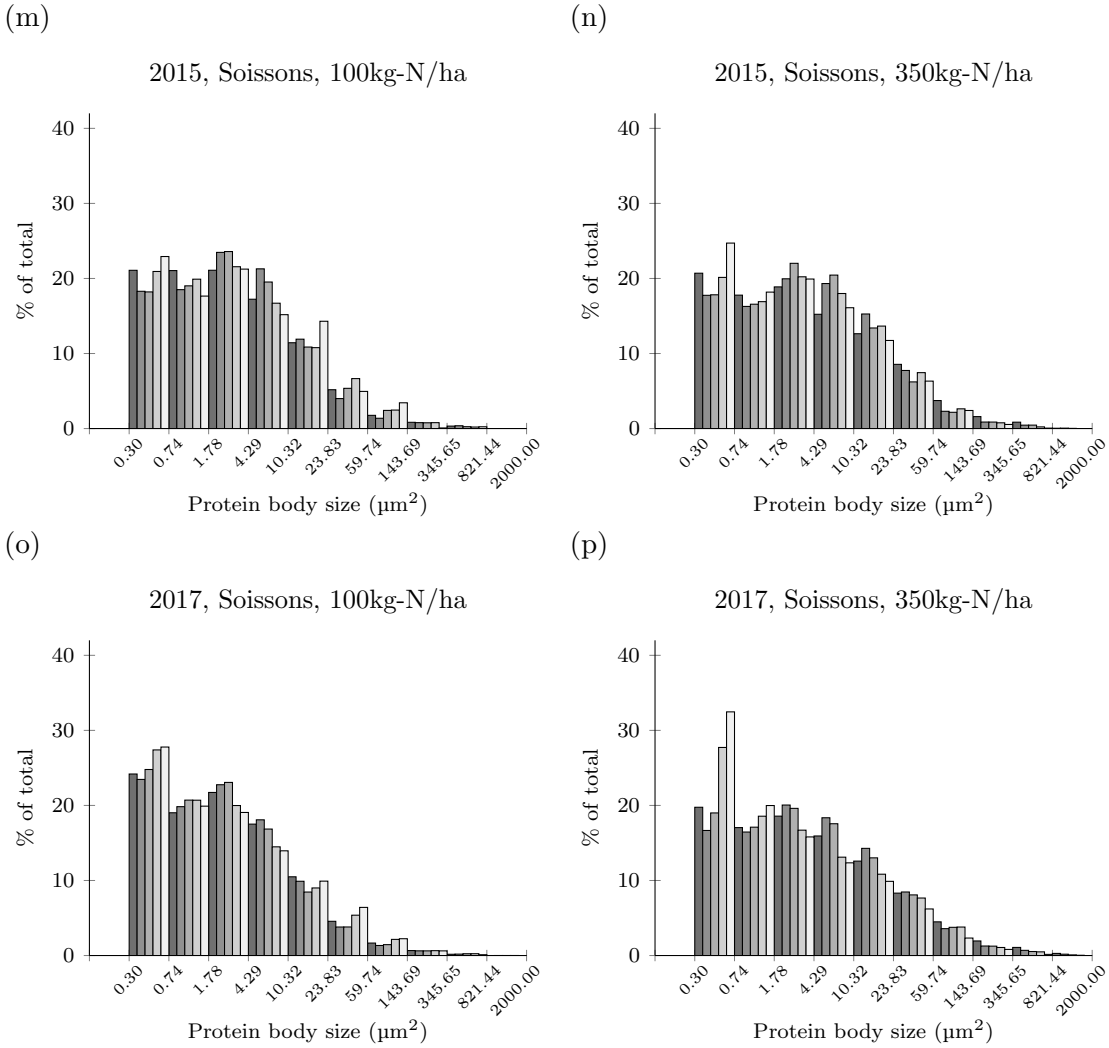


Figure 5.4: **Abundance of different sizes of protein bodies varies across the endosperm, with differential effects of nitrogen in different genotypes over different years**. Histograms showing the frequency of protein body size across the five endosperm zones, from outer in inner endosperm, represented by dark to light grey bars respectively. Each individual sub-figure shows the frequency of protein body sizes in each of the five zones, for a single combination of year of experiment, genotype, and nitrogen treatment: Cadenza in 2015 under (a) 100kg-N/ha and (b) 350kg-N/ha, and in 2017 under (c) 100kg-N/ha and (d) 350kg-N/ha; Hereward in 2015 under (e) 100kg-N/ha and (f) 350kg-N/ha, and in 2017 under (g) 100kg-N/ha and (h) 350kg-N/ha; Istabraq in 2015 under (i) 100kg-N/ha and (j) 350kg-N/ha, and in 2017 under (k) 100kg-N/ha and (l) 350kg-N/ha; Soissons in 2015 under (m) 100kg-N/ha and (n) 350kg-N/ha, and in 2017 under (o) 100kg-N/ha and (p) 350kg-N/ha

.

figure 5.3.

Histograms showing the abundance of different sizes of protein bodies over the five endosperm zones are presented in figure 5.4 for each combination of genotype and nitrogen treatment over the two years of the experiment. Again, the information presented in these histograms is discussed in the discussion at the end of this chapter (section 5.4.2).

5.4. **Discussion**

5.4.1. Controlled-environment experiment

The analysis of light-microscopy images of grain sampled from the controlled-environment shows that temperature and nitrogen input interact to determine a gradient in both the total protein concentration, and the size-distribution of protein bodies. Grain was sampled at two timepoints during mid to late grain-filling, with significant differences observed between these two timepoints. For the purpose of extrapolating the observations made in the grain protein concentration analysis to grain at maturity, the observations made at the later timepoint are most relevant, since they were made closer to the cessation of grain protein accumulation.

The protein concentration analysis from the controlled-environment experiment identified a linear decreasing gradient in protein concentration in all treatment combinations, with protein concentration decreasing towards the centre of the endosperm, confirming previous reports (Cobb 1905; Kent 1966; Tosi *et al.* 2011). Furthermore, this analysis shows for the first time that increased temperature during grain-filling results in more protein concentrated in the outer endosperm cells, and that an interaction exists between temperature and nitrogen supply. Under the low nitrogen treatment, the effect of temperature was comparable between sampling timepoints, whilst under high nitrogen input the effect of temperature was greater later in grain-filling. Since the effects aren't diminishing through development (*i.e.* between the early and later timepoint), it is likely that the effects of temperature and nitrogen supply on the distribution of protein in the developing wheat endosperm are preserved, or even enhanced, by the time grain is fully mature. These flat or increasing trends add confidence in our ability to use these results from mid to late grain-filling to predict the protein concentration gradients present in mature grain. Therefore this study provides evidence that grain grown in hotter temperatures will have a higher proportion of protein removed during the milling of white flour than grain grown under more temperature conditions, and also that the protein concentration of different mill streams will change, which must be accounted for in the blending of flour. Additionally, it shows how nitrogen supply interacts with temperature during grain-filling, causing greater changes in protein distribution when temperature are higher. In making comparisons with the field-grown wheat sampled from the WGIN diversity field trial experiment, it is interesting to note that the gradients observed in field-grown grain were more comparable to the gradients from plants subjected to the elevated (28°C) temperature treatment than the control (20°C) treatment when grown in controlled-

environment. This is unexpected, since the average day and night temperatures experienced during grain-filling under the high temperature treatment in the controlled environment experiment were considerably higher (23.7°C) than those recorded in either year of the field experiment (15.6°C in 2015, 17.36°C in 2017). This observation likely represents a pot effect, whereby the root systems of indoor-grown plants are less efficient and accumulate less protein in the grain, which results in a decrease in the gradients in protein concentration across the endosperm. This assumption is supported by the grain protein content measurements taken on mature grain (presented in chapter 6), with plants grown under the high (350kg-N/ha) nitrogen treatment in the field recording an average grain protein content of 13.63%, compared to a grain protein content of 9.32% in plants grown under the control (20°C) temperature treatment with high nitrogen input in controlled environment experiment.

The size-distribution analysis of protein bodies within the endosperm of grain grown in the controlled environment experiment describes the effect that temperature and nitrogen have on the average protein body size relative to its distance from the aleurone layer. A general linear trend for protein body size to decrease towards the centre of the endosperm was observed, with a greater gradient in protein body size-distribution recorded under higher levels of nitrogen input. Between the early and mid to late grain-filling sampling timepoints there was a reversal in the observed response to elevated temperature: earlier in grain-filling, protein bodies were generally larger under control temperature than elevated temperature, whilst later in grain-filling the opposite was true. Furthermore, at the later sampling timepoint the gradient in protein body size-distribution was less pronounced, and was practically non-existent under the low nitrogen treatment. The latter observation suggests that unlike the observations made in the protein concentration analysis, the gradients in protein body size-distribution decrease as grain-filling progresses. However, in this case predicting the state of mature grain from observations made on developing grain was never an objective, since the protein bodies fuse to form a protein matrix at maturity. Therefore the analysis of protein body size-distribution is more useful as a means of observing the initiation and growth of protein bodies during development, rather than for predicting the quality of mature grain.

The histograms presented in figure 5.2 show the results of the protein body size-distribution analysis (presented in figure 5.2) in more detail, with the abundance of protein bodies in each of the five endosperm zones sorted according to size. These figures illustrate not only the effects of temperature and nitrogen input over the two sampling timepoints, but also the general distribution of data. The requirement for the log-transformation used in the REML analysis is clear, since the majority (approximately 60–70%) of the protein bodies present were very small (with an area less than 4.29µm^2). The general trend of protein bodies increasing in size between the two sampling timepoints can be seen in the histograms, with a general "flattening" of the histograms at the later timepoint. This is to be expected, and shows that the analysis is detecting the growth of protein bodies over time. The histograms also provide some evidence for the mechanism by which gradients in the protein body size-distribution

come about. However, it should be noted that this analysis is based on the *relative* abundance of protein bodies according to size, and no absolute information is available.

As noted above, the greatest gradients in protein body size-distribution were identified at the earlier sampling timepoint. However, whilst the overall gradients are greater, the differences in the steepness of these gradients between temperature and nitrogen treatments are small (shown in figures 5.2a and 5.2b). Of the four combinations of temperature and nitrogen treatment, grain grown under elevated temperature with high nitrogen supply stands out as having the steepest negative gradient, whilst the other three treatment combinations show similar gradients in protein body size-distribution. This is mirrored in the histograms, with similar size-distributions observed for each treatment combination (see figures 5.3a, 5.3c, and 5.3e). These three histograms can be summarised by saying that the abundance of the smallest protein bodies was lowest in the far outside and inside (zones 1–2 and 4–5 respectively) of the endosperm, and that the abundance of medium protein bodies was greatest in these areas, whilst there was minimal differences in the abundance of the largest protein bodies across the endosperm. These differences in protein body sizes around the midpoint between inner and outer endosperm (zones 2–4) can be seen in the results of the REML analysis (figures 5.2a and 5.2b), whereby the plotted means appear to curve. This effect was only observed in the results from the early sampling timepoint, and after investigation the linear trend identified by the REML analysis was found to be the most suitable interpretation of the data. In contrast to the other three treatment combinations, for the elevated temperature, high nitrogen treatment combination, there is a higher abundance of the smallest (0.30–0.74µm^2) protein bodies in the inner endosperm, whilst a pattern similar to the other treatment combinations is seen in medium- and large-sized protein bodies. If the assumption is made that the smallest protein bodies represent recently initiated bodies, this suggests that under these conditions there are relatively more newly-formed protein bodies in the inner endosperm than the rest of the grain. However, due to the fact that only the relative number of protein bodies are analysed, this could indicate either an increase in the number of newly formed protein bodies, *i.e.* a stimulating effect on protein body initiation, or a lack of large protein bodies, indicating a delay in the onset of protein body initiation in this area of the grain.

In contrast to the results from the early sampling timepoint, at the later timepoint the gradients in protein body size-distribution are less pronounced, and almost non-existent under the low nitrogen treatment (see figures 5.2c and 5.2d). The histograms for this data are shown in figures 5.3b, 5.3d, 5.3f, and 5.2h, and show a number of trends of interest. The clearest general trend shown across all treatments at the mid to late grain-filling sampling timepoint is seen in the distribution of the largest protein bodies: the outer endosperm contains the highest proportion of large protein bodies, with the proportion decreasing towards the centre of the endosperm. However, whilst this trend is shown across all treatment combinations, it is more pronounced under both the control-temperature treatment and the high nitrogen treatment, which contributes to the increased gradient in protein body size-distribution under

the combined control-temperature and high nitrogen treatments. The histograms for the low nitrogen results are particularly interesting, since a minimal gradient in protein body size-distribution is observed in the REML analysis (figure 5.2c), but the histograms (figures 5.3b and 5.3f) show that there are a number of differences in the size-distribution of protein bodies that essentially cancel each other out when observing the means alone. For example, under the control temperature and low-nitrogen treatment combination (figure 5.3b), there is an increase in the abundance of medium-sized protein bodies towards the central endosperm, but a decrease in the abundance of small and large protein bodies. When compared to the histogram for the respective high-nitrogen treatment, the main difference is an increase in the abundance of small protein bodies in the inner endosperm. This difference, although subtle, is enough to create the gradient shown in figure 5.3d. However, although small, the change in the abundance of the smallest protein bodies is important, since it represents a relative increase in newly produced protein bodies. Under elevated temperature, the effect of nitrogen input on the protein body size-distribution analysis was reduced. However, the effects described above can still be observed in the histograms (figures 5.3f and 5.2h), with an increase in the abundance of small protein bodies linked with an increase in nitrogen supply.

5.4.2. WGIN diversity field trial experiment

The results from the protein concentration gradient analysis of grain sampled from the WGIN diversity field trial experiment found a significant interaction between the effects of nitrogen, year of experiment, and genotype. As in the controlled-environment experiment, all treatment combinations showed a linear gradient in total protein concentration that decreased towards the centre of the grain, with increased nitrogen input increasing this gradient. Whilst the general response to nitrogen is characterised by an increase in the protein concentration gradient, the magnitude of this effect varied between genotypes and years. The two years sampled as part of this experiment, 2015 and 2017, represent comparable years in terms of rainfall and sunlight, but cooler temperatures were experienced during grain-filling in 2015 (see chapter 3). Therefore in a limited sense 2015 could be viewed as a control year, and 2017 as an "elevated temperature" year. However, the difference in average temperature during grain-filling (day and night inclusive) between 2015 and 2017 was only 0.82°C, compared to a difference of 5.33°C between the two temperature treatments used in the controlled-environment experiment. This small difference in average temperature may go some way to explain the minimal differences in grain protein concentration gradient between years: whilst the overall protein concentration (shown by the position of the trend line in the direction of the y axis in the figures) is generally higher in 2017 compared to 2015, the gradient in protein concentration (*i.e.* the angle of the trend lines) shows minimal difference between years. Whilst the overall effect of year was minimal, small differences in the response to nitrogen fertiliser were shown by each genotype in each year. Hereward, Istabraq, and Soissons all showed a slightly increased response to nitrogen in 2017 compared to 2015, whilst Cadenza showed a slightly reduced response in 2017. Furthermore, both the response to nitrogen, and

the change in response to nitrogen over the two years of the experiment shown by Hereward, Istabraq, and Soissons was almost identical. With regards to Cadenza, in 2017 it only shows a slightly greater response to nitrogen compared to the other genotypes. However in 2015, the response to nitrogen shown by Cadenza was more than double the response shown by the other genotypes. Therefore it is likely that the large nitrogen response of Cadenza in 2015 resulted in the interaction between genotype, nitrogen, and year of experiment that was identified as part of the REML analysis. More generally, there is a clear divide among the protein concentration gradients observed in each genotype, with the high-protein bread-making varieties, Cadenza, Hereward, and Soissons, having the greatest gradients, and the low-protein feed wheat Istabraq the smallest.

With regards to the protein body size-distribution analysis, the results from the WGIN diversity field trial experiment are somewhat complementary to those collected from the controlled-environment experiment, and show that increased nitrogen input also increases the gradient in protein body size in field-grown wheat. Additionally, this dataset shows how different genotypes show different responses to nitrogen input. Of the four genotypes, Istabraq stands out for its lack of response to nitrogen input, with no clear response in either 2015 or 2017. Cadenza and Hereward both exhibited the greatest responses to nitrogen, with Cadenza showing the most consistent result over the two years of the experiment, whilst Hereward's response was reduced in 2017. In comparison to nitrogen, the effect of year appears to be minimal. The only genotypes that shows a clear response to year of experiment are Soissons and to a lesser extent, Istabraq. However, these genotypes responded differently: Istabraq had greater gradients in protein body size-distribution in 2015, whilst Soissons achieved the greatest change between years, with an increase in the negative gradient under the 350kg-N/ha treatment, and a reversal from positive to negative gradient between 2015 and 2017 under the 100kg-N/ha treatment. However, the amount of missing data as part of this analysis should be noted, since only one biological replication was available for analysis in both Istabraq and Soisson in 2015, 100kg-N/ha treatment, these findings are presented with a reduced level of confidence compared to those supported by higher levels of replication.

The histograms exploring the size-distribution of protein bodies from the WGIN field trial experiment (presented in figure 5.4) show that the data is skewed towards an abundance of small protein bodies (as discussed in section 5.4.1). However, in comparison to the analysis conducted on the controlled-environment experiment, the histograms are generally flatter, with a greater abundance of larger protein bodies. This is likely due to the fact that the accumulated thermal time between anthesis and sampling in the field experiment was considerably greater than that in the controlled-environment experiment, and as a result the grain is likely to be closer to physiological maturity. Another general trend in the data from the field experiment that is different to the controlled-environment experiment results is the abundance of the smallest protein bodies (up to 0.74μm^2) in the central endosperm under the high nitrogen treatment (350kg-N/ha) of all genotypes used in the study. As discussed in section 5.4.1,

this increase in the proportion of the smallest protein bodies could represent an increase in the number of newly formed protein bodies, or a *decrease* in the number of larger, more developed protein bodies. Due to difference between the field and controlled-environment it is difficult to attribute a cause to this observation, as it could be due to any number of differences between field and controlled-environment experiments. However, nitrogen availability is clearly involved in this increase in the proportion of small or newly formed protein bodies in the central endosperm, either through amount applied or amount available to the plant, which confirms a similar, albeit smaller, response to nitrogen input in the controlled-environment experiment.

The clearest difference between the protein body size-distribution results from the field experiment and the controlled-environment experiment (figures 5.3 and 5.2 respectively) is the presence of positive (increasing) gradients for Cadenza and Soissons in 2015 (based on a single biological replication) under the 100kg-N/ha treatment. The histograms (see figures 5.4a, 5.4c, and 5.4i) show that these increasing gradients are largely caused by an increase in the abundance of large protein bodies, but not a decrease in the proportion of the smallest (less than 0.74µm^2) protein bodies.

The histograms representing the protein body size-distribution data also show some of the differences between the four genotypes studied in this experiment. Firstly, when comparing the overall shape of the histograms, ignoring the distribution over endosperm zones, the high-protein bread-making wheats Cadenza and Hereward generally have a more linear distribution of protein body sizes, *i.e.* fewer medium-size bodies. Istabraq and Soissons, however, have a relatively greater proportion of these medium-sized protein bodies, which can be observed as a peak that is shifted down the x axis. Additionally, when comparing between these two sets of genotypes, there are more of the largest protein bodies in Cadenza and Hereward. By combining these observations it can be stated that these high-protein varieties contain more small, newly initiated protein bodies, as well as more large, well established protein bodies, which results in an overall increase in grain protein concentration.

Chapter 6: Grain protein composition results

6.1. Introduction

Wheat is defined by the properties of its flour. When wetted, wheat flour forms a viscoelastic dough that is both elastic and extensible (Lásztity 1996). These desirable dough characteristics are largely determined by the protein content and composition of the mature grain, with different protein compositions suited to different end-uses. Depending on the storage proteins present in the mature grain, wheat can be used for the production of a wide range of foods, from pasta, to biscuits, to bread. Therefore understanding the factors that influence the accumulation of protein in the wheat grain is of great importance.

The endosperm of the mature wheat grain contains 8–20% protein (Davis *et al.* 1981), and of this protein approximately 80% is gluten storage protein (Peña 2002). Both the total amount of gluten protein present in the grain, and the composition of this gluten protein determine the quality of the dough produced (El Haddad *et al.* 1995; Sapirstein *et al.* 1998). The gluten storage proteins can broadly be split into two groups: the polymeric glutenins, and the monomeric gliadins. The glutenins consist of the HMW and LMW subunits, and confer strength and elasticity to the dough, whilst gliadins consist of alpha-, beta-, gamma-, and omega-gliadin proteins, and determine the viscosity or flow of the dough (Uthayakumaran *et al.* 2000).

The accumulation of storage proteins during grain-filling is a dynamic process, and is affected by climate, nitrogen fertiliser input, and genotypic variation (Bergman *et al.* 1998; Hurkman *et al.* 2013; Chope *et al.* 2014). To investigate how these factors affect protein quality in the context of UK bread-making, both protein content and composition was measured on grain grown both in the field, and under controlled-environment conditions. To investigate the effects of climate, nitrogen input, and genotype in the field, samples were taken from three years of the WGIN (www.wgin.org.uk) diversity experiment, a long-term field trial experiment in which multiple wheat genotypes are grown under different nitrogen inputs. A controlled-environment experiment was undertaken to identify the effects that elevated temperature during grain-filling and the level of nitrogen input prior to anthesis have on protein accumulation and quality.

Grain protein concentration/content was determined through nitrogen content analysis using the Dumas method, and protein composition was measured by both SDS-PAGE and SE-HPLC, from which the gluten protein content was calculated. Protein composition measurements were taken to predict the bread-making quality of the sampled grain.

6.2. Controlled-environment experiment

The combined effect of temperature and nitrogen supply were investigated in the controlled-environment experiment, which used a control (20°C) and an elevated (28°C) post-anthesis

daytime temperature treatment, and a high and low nitrogen treatment to investigate the effect these factors have on protein accumulation during grain-filling. Grain protein concentration was measured both during grain-filling, and at maturity. From this data, the protein *content* of the mature grain was calculated, to determine whether any changes in protein concentration were due to changes in the amount of protein accumulated, or due to a dilution effect caused by other cellular components such as starch. The composition of grain protein was measured at maturity by SDS-PAGE and SE-HPLC to provide estimations of the bread-making quality of the grain, since the grain samples collected were too small to bake test loaves. SE-HPLC data was used to calculate the ratio between HMW- and LMW-glutenins (F1/F2), and the ratio between HMW-glutenins and gliadins ((F3+F4)/F1), both of which are useful predictors of bread-making quality. Protein composition was also measured by SDS-PAGE throughout grain-filling, allowing for comparisons to be made between the relative amount of the individual gluten protein subunits present. Such comparisons were possible since all samples were from the same genotype (Cadenza), and therefore contained the same protein subunits.

6.2.1. Protein concentration and content

Nitrogen content was measured on grain collected both at five timepoints during grain-filling, and again at maturity. For grain collected at maturity, the *content* of protein per grain was calculated from nitrogen content data and TGW measurements. Nitrogen content data was converted to protein using a factor of 5.7 (Sosulski *et al.* 1990), and is presented as such throughout.

The analysis of protein concentration data during grain-filling found a two-way interaction between post-anthesis temperature treatment and timepoint of measurement ($F_{5,60}$=15.46, P<0.001, SED=0.441 on 21.88 DF). This interaction describes how protein concentration is continuously decreasing under control temperatures, but decreases during early grain-filling before increasing again under elevated temperature, *i.e.* the accumulation pattern of protein over time is different between the two temperature treatments. In addition to this two-way interaction, there was also a two-way interaction between temperature and nitrogen treatment ($F_{1,6}$=7.59, P=0.033, SED=0.351 on 9.02 DF), which did not include the timepoint at which the measurement was taken, and so was a consistent effect throughout grain-filling. This interaction describes how the high nitrogen treatment resulted in an average increase of 9% under control temperatures, and an increase of 14% under elevated temperature. Both of the two-way interactions can be observed in the full dataset, presented in figure 6.1.

For protein concentration measured at maturity, a two-way interaction between temperature and nitrogen treatments was found ($F_{1,6}$=12.27, P=0.013, SED=0.265 on 10.86 DF) (figure 6.2): at maturity, grain protein concentration was significantly higher under the high-temperature treatment, and increased nitrogen also increased protein concentration, albeit to a lesser extent. The effect of the high-nitrogen treatment on grain protein concentration was greater when temperatures were elevated, with an average increase of 15% compared to an

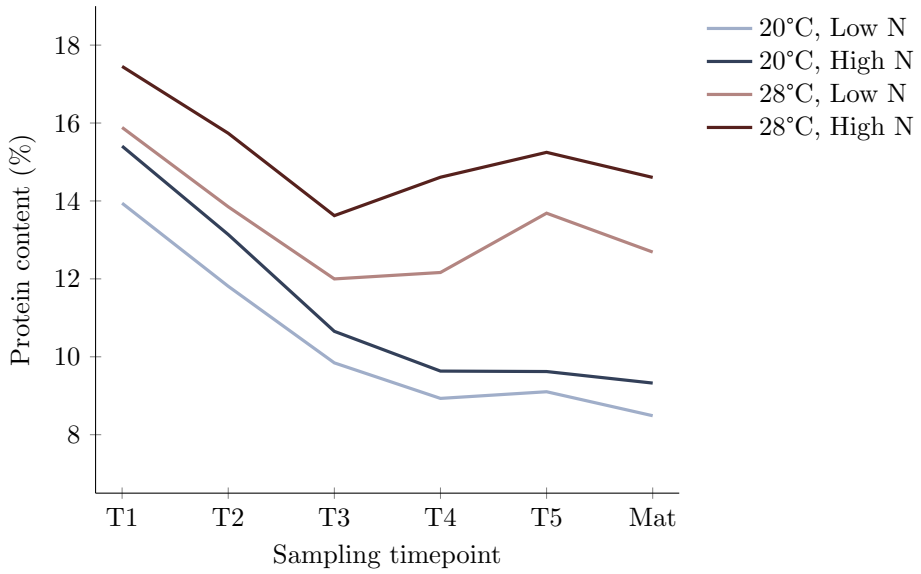


Figure 6.1: **Elevated post-anthesis temperature and increased nitrogen input interact to increase grain protein concentration throughout grain-filling.** Grain protein concentration data from the controlled-environment experiment, calculated as concentration of nitrogen multiplied by 5.7. Individual treatments of control temperature, low nitrogen (light blue) and high nitrogen (dark blue), and elevated temperature low nitrogen (light red) and high nitrogen (dark red) are shown.

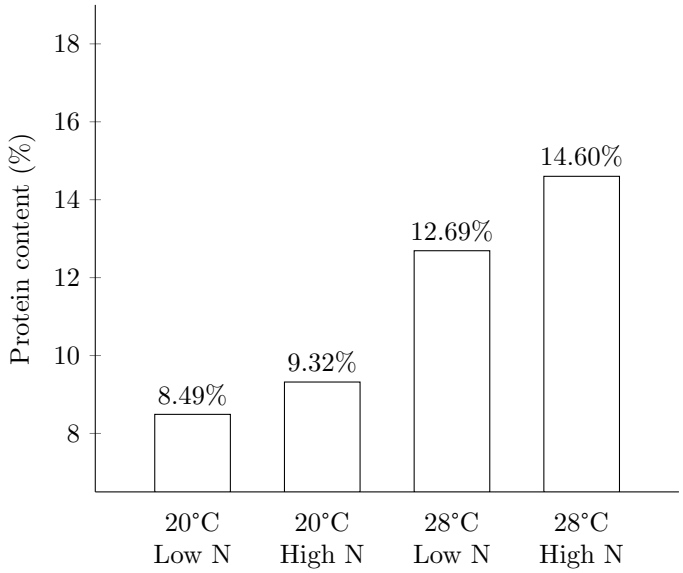


Figure 6.2: **Elevated post-anthesis temperature and increased nitrogen input interact to increase the protein concentration of mature grain.** Mature grain protein concentration data from the controlled-environment experiment, grouped by individual treatments of control temperature (20°C), low-, and high-nitrogen, and elevated temperature (28°C), low-, and high-nitrogen. LSD (at the 5% level) of 0.53 for comparisons within the same temperature treatment, and 0.58 for all other comparisons.

.

Table 6.1: Full dataset for grain protein content at maturity from the controlled-environment experiment, presented per treatment combination, averaged over experimental blocks. Expressed as protein per grain.

Treatment	Protein content per grain (mg)
20°C, low-nitrogen	3.39
20°C, high-nitrogen	4.01
28°C, low-nitrogen	3.59
28°C, high-nitrogen	4.32

increase of 10% under the control temperature treatment.

Protein content calculated from measurements taken from mature grain increased 16% from 3.59mg per grain to 4.16mg under the high nitrogen input treatment ($F_{1,6}$=39.02, P<0.001, SED=0.263 on 6 DF) (figure 6.3). Protein content was not found to be significantly different between post-anthesis temperature treatments ($F_{1,5}$=2.45, P=0.178), and no significant interaction between temperature and nitrogen treatments was found ($F_{1,6}$=0.27, P=0.622). The full dataset is presented in table 6.1.

6.2.2. SDS-PAGE Results

Sodium-dodecyl-sulphate polyacrylamide gel electrophoresis analysis was completed on protein extracted during grain-filling, and again at maturity, to describe how individual gluten proteins were affected by both elevated post-anthesis temperature and limited nitrogen supply. The area under the curve was measured for each protein band in the intensity graph produced during the analysis of each gel, and the size of each band was then expressed as a percentage of the total protein detected. This approach was taken to account for differences in the efficiency

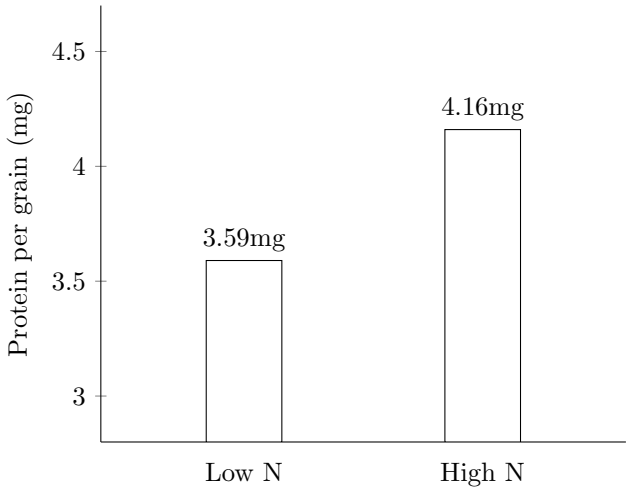


Figure 6.3: **Increased nitrogen supply during vegetative development increases the protein content of mature grain**. Protein content of the mature grain from the controlled-environment experiment, grouped by nitrogen treatment, and averaged across temperature treatments. LSD (at the 5% level) of 0.26.

of the protein extraction, and of the staining, de-staining, and imaging of the gel. Additionally, it allows for the protein composition to be analysed without it being compounded by any differences in total protein content.

As an exploratory technique, canonical variate analysis was conducted on the SDS-PAGE data, and aims to distinguish the effect of each treatment combination on the protein composition profile of each sample as a whole. Figure 6.4 shows the results of this analysis, and predicts that the largest factor influencing protein composition is time (represented by CV1), with the second largest impact coming from the temperature treatment (CV2), and with nitrogen input showing no clear differentiation in either CV1 or CV2. The analysis suggests that the effect of temperature treatment increases as grain-filling progresses, but also that the protein composition from timepoint T4 to maturity is relatively static, suggesting that protein composition was largely determined at this stage. Figure 6.5 shows the vector loadings for each individual protein in the CV analysis, and illustrates which proteins had the largest impact on the separation of treatments shown in figure 6.4. This allows the identification of proteins which are more susceptible to increased post-anthesis temperature, or are differentially

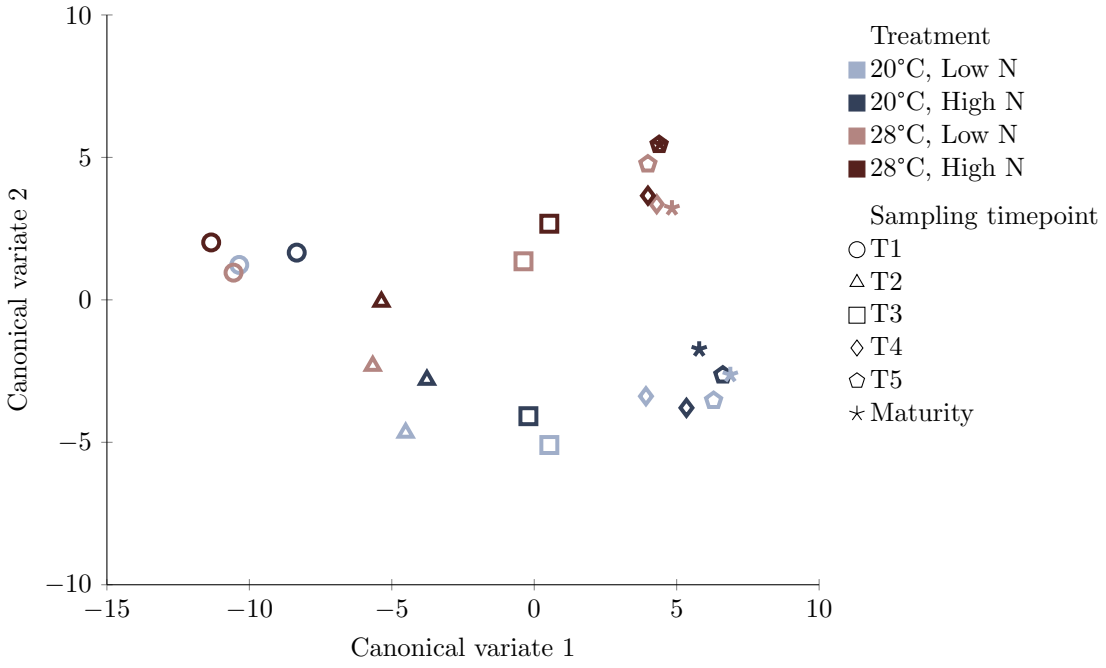


Figure 6.4: **Grain protein composition is differentially affected by temperature and time during grain-filling.** CV analysis of SDS-PAGE protein composition data from the controlled-experiment. Analysis separates the individual combinations of temperature, nitrogen, and sampling timepoint based on the overall protein composition. CV1 primarily separates the effect of time, whilst CV2 separates the two post-anthesis temperature treatments. Nitrogen treatment shows less separation, that isn't clearly represented by either CV1 or CV2. CV1 and CV2 account for 67.79% and 21.72% of the observed variation respectively.

.

Table 6.2: Results of the REML analyses of SDS-PAGE data throughout grain-filling from the post-anthesis controlled environment experiment. Table shows the significant effects of all treatment combinations on individual protein bands from the SDS-PAGE analysis. Significant results to be interpreted are shown in bold. The single effect of sampling timepoint was ignored, since it doesn't relate to the experimental treatments applied as part of the experiment.

HMW-glutenins

Fixed term	P1	P2	P3	P4
Temperature	**<0.001**	**<0.001**	0.003	0.368
Nitrogen	0.283	0.205	0.377	0.978
Timepoint	<0.001	<0.001	<0.001	<0.001
Temperature.Nitrogen	0.939	0.732	0.080	0.392
Temperature.Timepoint	0.746	0.281	**<0.001**	0.684
Nitrogen.Timepoint	0.185	0.127	0.122	0.407
Temperature.Nitrogen.Timepoint	0.375	0.674	0.406	0.120

Omega-gliadins

Fixed term	P5	P6	P7
Temperature	0.025	<0.001	<0.001
Nitrogen	0.002	<0.001	0.053
Timepoint	<0.001	<0.001	<0.001
Temperature.Nitrogen	**0.002**	0.069	**0.040**
Temperature.Timepoint	**0.006**	**<0.001**	**0.009**
Nitrogen.Timepoint	0.498	**0.009**	**0.032**
Temperature.Nitrogen.Timepoint	0.296	0.885	0.705

LMW-glutenins

Fixed term	P8	P9	P10
Temperature	<0.001	0.717	0.331
Nitrogen	**<0.001**	0.728	0.011
Timepoint	<0.001	<0.001	0.003
Temperature.Nitrogen	0.498	0.066	**0.019**
Temperature.Timepoint	**<0.001**	0.842	0.919
Nitrogen.Timepoint	0.133	0.091	0.419
Temperature.Nitrogen.Timepoint	0.483	0.716	0.910

Alpha-, beta-, and gamma-gliadins

Fixed term	P11	P12	P13
Temperature	0.131	<0.001	**<0.001**
Nitrogen	0.631	0.671	0.918
Timepoint	<0.001	<0.001	<0.001
Temperature.Nitrogen	0.332	0.129	0.398
Temperature.Timepoint	**0.002**	**0.012**	0.694
Nitrogen.Timepoint	0.484	0.399	0.162
Temperature.Nitrogen.Timepoint	0.595	0.277	0.842

accumulated at different stages of development. From this analysis, protein subunits P3 and P4 (both HMW-GS) stand out as being susceptible to post-anthesis temperature, as they both have strong influence on the separation shown by CV2.

REML analysis was used to describe how each individual protein subunit was affected by the temperature and nitrogen treatments over the period of grain-filling (data presented in full in figure 6.6). Separate analyses were completed for each detected protein, and the P values from these analyses are presented in table 6.2. All of the detected proteins showed a response to sampling timepoint, showing that the relative level of each protein significantly changes over the period of grain-filling. However, as a single effect this response doesn't relate to the experimental treatments of nitrogen input and post-anthesis temperature, it wont be discussed.

Of the four HMW-glutenin proteins present in Cadenza, three showed an effect from temperature: protein P1 (F=43.38, P<0.001, SED=0.1216 on 66.2 DF) and P2 (F=25.08, P<0.001, SED=0.1047 on 66.4 DF) both saw an increase in relative concentration under elevated temperature, whilst protein P3 was affected by a two-way interaction between timepoint and temperature (F=6.26, P<0.001, SED=0.1797 on 64.6 DF), with relative concentration

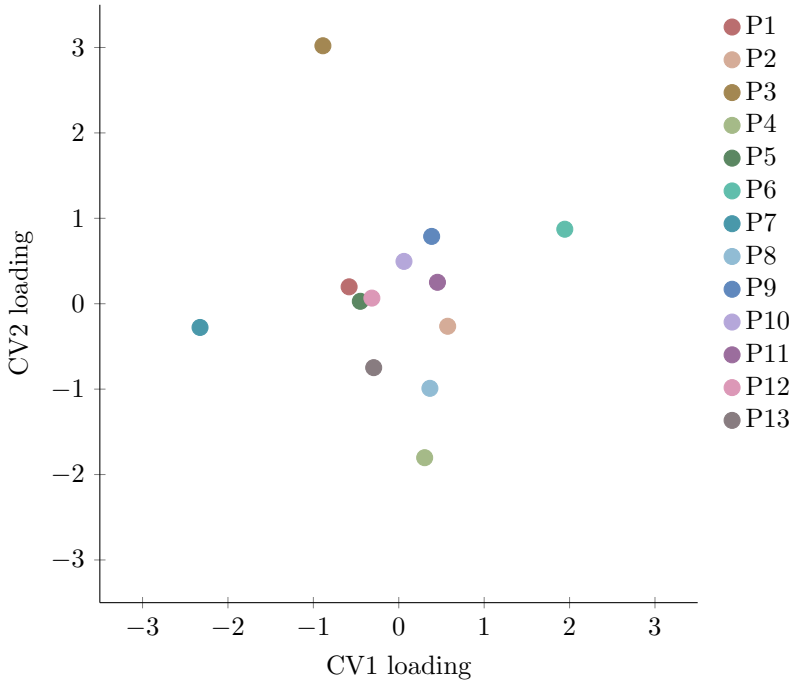


Figure 6.5: **HMW-glutenin subunits P3 and P4 both show strong influence on the separation of temperature treatments by CV analysis.** Vector loadings for CV analysis of SDS-PAGE data from the controlled-environment experiment. Loadings indicate the influence each protein has on the CV analysis presented in figure 6.4. HMW-glutenins P3 and P4 both show strong influence on CV2, which represents the separation of temperature treatments.

.

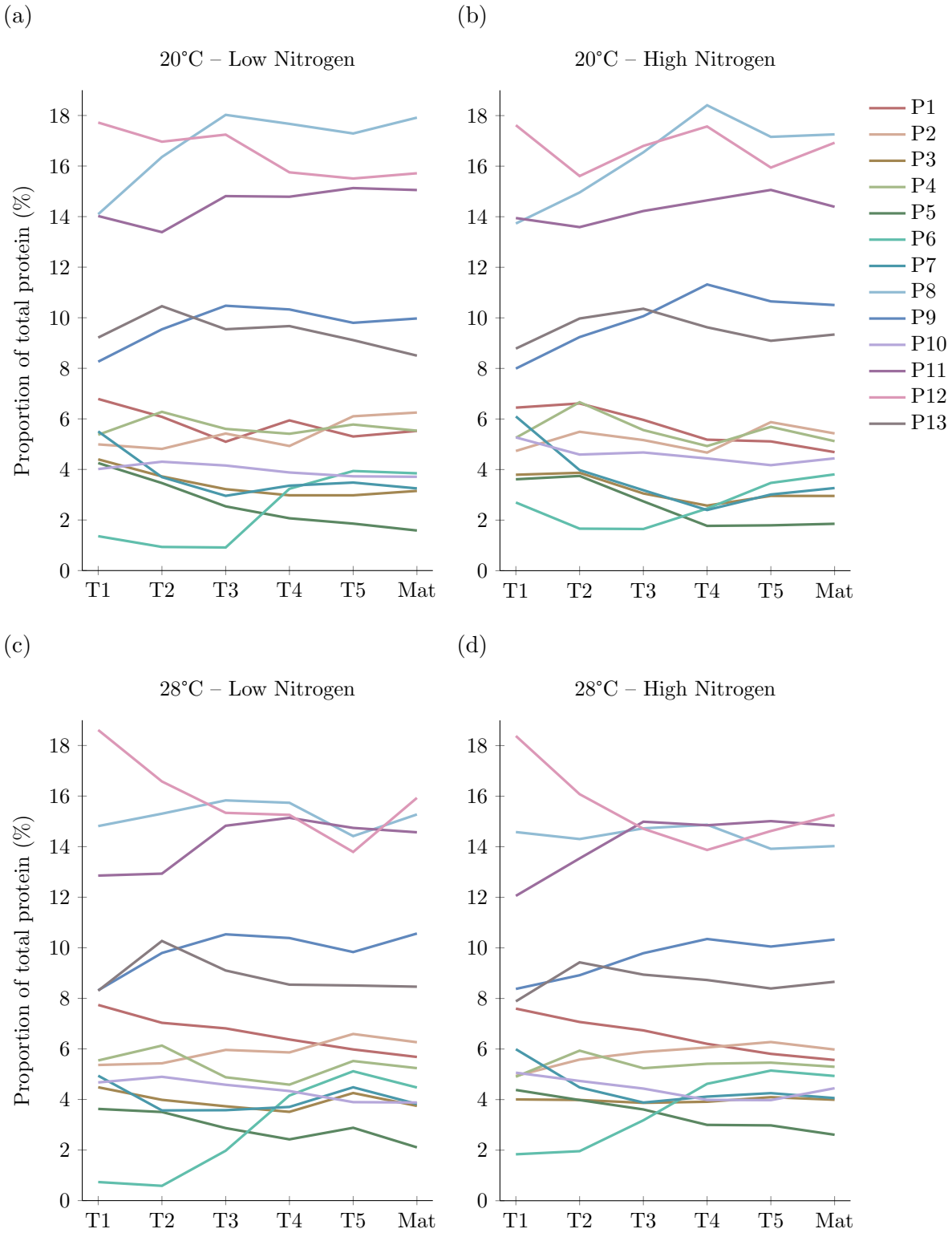


Figure 6.6: **Storage protein subunits are differentially accumulated during development under different temperature and nitrogen regimes.** Grain protein subunit abundance detected by SDS-PAGE throughout grain development from the controlled-environment experiment, grouped by treatment combination: (a) control temperature, low nitrogen; (b) control temperature, high nitrogen; (c) elevated temperature, low nitrogen; (d) elevated temperature, high nitrogen.

increasing in the elevated temperature treatment, and this effect increasing as grain-filling progresses.

All three of the omega-gliadins analysed showed multiple two-way interactions. The concentration of protein P5 was increased by high nitrogen input, but only in the elevated temperature treatment (F=10.63, P=0.002, SED=0.1640 on 64.8 DF), and whilst the concentration of P5 generally decreased during grain-filling, the overall concentration was higher, and the drop in concentration was less under higher temperatures (F=3.66, P=0.006, SED=0.2303 on 64.4 DF). Protein P6 concentration generally increased during grain-filling, but increased at a greater rate under elevated post-anthesis temperatures, reaching a higher final level (F=6.49, P<0.001, SED=0.3559 on 68 DF); whilst an interaction between time and nitrogen was also significant (F=3.34, P=0.009, SED=0.3555 on 67.1 DF), the effect was only significant at the early two sampling timepoints, where increased nitrogen increased the proportion of P6. However this effect disappeared as grain-filling progressed. For the relative concentration of omega-gliadin P7, three two-way interactions were significant: over grain-filling, the concentration of protein P7 was higher under the high temperature treatment when compared with the control temperature treatment, with this difference increasing towards maturity (F=3.33, P=0.009, SED=0.2862 on 68.4 DF); under the high nitrogen treatment, the proportion of P7 was initially increased compared to the low nitrogen treatment, but by the end of grain-filling there wasn't a significant difference between the two nitrogen treatments (F=2.62, P=0.032, SED=0.2859 on 67.4 DF); and there was a weak interaction between temperature and nitrogen treatments, with the proportion of P7 increasing with increased nitrogen input, but only when subjected to elevated temperatures (F=4.38, P=0.040, SED=0.1650 on 69 DF).

With regards to the LMW-glutenins detected by SDS-PAGE, two proteins showed significant effects from factors other than time alone. Protein P8 increased in concentration through grain-filling, and did so at a greater rate when exposed to higher temperatures (F=15.65, P<0.001, SED=0.4087, DF not available). A single effect of nitrogen on the proportion of P8 was also identified, with higher levels of nitrogen application decreasing the relative concentration of P8 (F=22.88, P<0.001, SED=0.1558, DF not available). A significant two-way interaction between temperature and nitrogen was found for protein P10, with the proportion of P10 increasing due to high nitrogen input under control temperature, but *decreasing* under high temperatures (F=5.82, P=0.019, SED=0.1839 on 68.2 DF).

The alpha-, beta-, and gamma-gliadins analysed all showed a significant response to temperature. The REML analysis reported a two-way interaction between temperature and sampling timepoint on the proportion of protein P11 (F=4.30, P=0.002, SED=0.3381 on 69.2 DF), however comparison of the means with the LSD at the 5% level reveals that the difference between the two temperature treatments is only significant at the earliest sampling timepoint. The single effect of temperature was not found to be significant (F=2.33, P=0.131). The same two-way interaction between temperature and time was also significant for protein P12

(F=3.19, P=0.012, SED=0.3381 on 66.9 DF), but again, inspection of the means reveals that the interaction isn't particularly interesting. However, P12 also showed a significant effect from temperature alone (F=13.12, P<0.001, SED=0.2433 on 66.1 DF), with the level of P12 decreasing under elevated temperature. Finally, P13 also showed a single effect of temperature (F=24.38, P<0.001, SED=0.1416 on 66.5 DF), with elevated temperature again decreasing the proportion of P13.

A separate REML analysis was completed for the SDS-PAGE data from only mature grain, to identify any effect the nitrogen and temperature treatments may have on the protein composition of the grain at harvest. The results of these analyses are presented in table 6.3, with the interpreted P values shown in bold, and the full table of means is presented in table 6.4.

Significant two-way interactions between the temperature and nitrogen treatments were found for two of the HMW-glutenin proteins, P2 (F=6.61, P=0.033, SED=0.1467 on 8.1 DF) and P3 (F=7.06, P=0.021, SED=0.1175 on 12 DF). The proportion of protein P2 was primarily affected by nitrogen, and was reduced under high nitrogen input, however comparison of the means with the LSD at the 5% level shows that this difference was only significant under the control temperature treatment. Additionally, the high temperature treatment *increased* the proportion of protein P2, but only in the high nitrogen treatment. Whilst the interaction between temperature and nitrogen treatments for protein P3 was significant, further analysis revealed that this interaction didn't describe any significant differences when comparing means with the relevant LSD values. Therefore for protein P3, the single effect from temperature is interpreted instead (F=96.56, P<0.001, SED=0.0831 on 12 DF): higher temperature during grain-filling increased the proportion of protein P3.

The proportion of all three omega-gliadin proteins was increased by the elevated temperature treatment: protein P5 (F=12.45, P=0.012, SED=0.1806 on 6 DF), protein P6 (F=34.03, P<0.001, SED=0.1581 on 10.5 DF), and protein P7 (F=16.83, P=0.002, SED=0.1728 on 10.5 DF). The proportion of omega-gliadin P5 was also increased by the high nitrogen treatment (F=17.53, P=0.006, SED=0.0925 on 6 DF).

LMW-glutenin protein P8 was significantly reduced under the elevated temperature treatment (F=72.23, P<0.001, SED=0.3377 on 10.5 DF), and was also reduced, albeit to a lesser extent, under the high-nitrogen treatment (F=9.10, P=0.013, SED=0.3332 on 10.2 DF). A weak significant two-way interaction between temperature and nitrogen treatments was identified for protein P9 (F=6.24, P=0.031, SED=0.2599 on 10.1 DF). However, inspection of the means revealed that this interaction is describing the increased proportion of P9 due to high temperature under the low nitrogen treatment. Other responses were not found to be significant at the 5% level. Finally, the proportion of protein P10 was increased by the high nitrogen treatment (F=7.10, P=0.021, SED=0.2444 on 12 DF).

Of the three alpha-, beta-, and gamma-gliadin proteins, P11 and P12 showed a significant

Table 6.3: Results of the REML analyses of SDS-PAGE data at maturity from the post-anthesis controlled environment experiment. Table shows the significant effects of all treatment combinations on individual protein bands from the SDS-PAGE analysis. Significant results to be interpreted are shown in bold. Weakly significant interactions have been excluded when further analysis reveals that they do not describe any significant differences between means of interest.

HMW-glutenins

Fixed term	P1	P2	P3	P4
Temperature	0.141	0.023	**<0.001**	0.656
Nitrogen	0.060	<0.001	0.754	0.205
Temperature.Nitrogen	0.153	**0.033**	0.021	0.099

Omega-gliadins

Fixed term	P5	P6	P7
Temperature	**0.012**	**<0.001**	**0.002**
Nitrogen	**0.006**	0.163	0.376
Temperature.Nitrogen	0.255	0.106	0.429

LMW-glutenins

Fixed term	P8	P9	P10
Temperature	**<0.001**	0.275	0.725
Nitrogen	**0.013**	0.637	**0.021**
Temperature.Nitrogen	0.314	**0.031**	0.743

Alpha-, beta-, and gamma-gliadins

Fixed term	P11	P12	P13
Temperature	0.755	**0.034**	0.314
Nitrogen	0.519	0.375	0.151
Temperature.Nitrogen	0.036	0.029	0.350

response to the experimental treatments. Whilst the two-way interaction between temperature and nitrogen treatment was significant for both P11 (F=6.21, P=0.036, SED=0.2927 on 8.3 DF) and P12 (F=6.48, P=0.029, SED=0.5025 on 10.2 DF), analysis of the means reveal that this interaction isn't the best descriptor of the data in either case. For protein P11, analysis of the means reveals no significant differences at the 5% level, and for protein P12, the proportion is increased by the high nitrogen treatment, but only under the control temperature treatment. However, further analysis of the data suggests that the two-way interaction is largely caused by a single, potentially anomalous, datapoint, and that the single effect of temperature (F=5.95, P=0.034, SED=0.3590 on 10.4 DF) may be a better description of the data, with higher temperature generally decreasing the concentration of protein P12.

6.2.3. SE-HPLC Results

Size-exclusion high-performance liquid chromatography analysis was completed on protein extracts from mature grain to predict the effect that nitrogen supply and post-anthesis temperature have on the baking quality of mature grain. The relative abundance of five protein fractions representing HMW-glutenins (F1), LMW-glutenins (F2), omega-gliadins (F3), alpha-, beta-, and gamma-gliadins (F4), and albumin and globulins (F5) were measured. From these measurements gluten content (F1–F4 as a percentage of F1–F5), the ratios between different protein groups (F1/F2) and (F3+F4)/F1), and the overall protein composition is

Table 6.4: Full dataset for SDS-PAGE grain protein composition data from the controlled-environment experiment. Data is presented as the relative abundance of each individual protein subunit as a percentage of the total protein, and is grouped by protein group, with sub-totals for each group. Mean values for each treatment combination are presented, averaged over experimental blocks.

		Proportion of the total protein (%)			
Protein group	Protein	20°C, low-N	20°C, high-N	28°C, low-N	28°C, high-N
HMW-glutenins	P1	5.52%	4.67%	5.69%	5.55%
	P2	6.25%	5.43%	6.27%	5.97%
	P3	3.15%	2.96%	3.75%	3.99%
	P4	5.53%	5.12%	5.24%	5.30%
	sub-total	*20.45%*	*18.18%*	*20.95%*	*20.81%*
Omega-gliadins	P5	1.58%	1.86%	2.11%	2.61%
	P6	3.81%	3.78%	4.46%	4.98%
	P7	3.23%	3.25%	3.80%	4.10%
	sub-total	*8.62%*	*8.89%*	*10.37%*	*11.69%*
LMW-glutenins	P8	17.87%	17.21%	15.35%	13.99%
	P9	9.97%	10.50%	10.64%	10.26%
	P10	3.71%	4.44%	3.88%	4.45%
	sub-total	*31.55%*	*32.15%*	*29.87%*	*28.70%*
Alpha-, beta-, and gamma-gliadins	P11	15.06%	14.42%	14.50%	14.86%
	P12	15.81%	17.03%	15.84%	15.25%
	P13	8.50%	9.34%	8.46%	8.66%
	sub-total	*39.37%*	*40.79%*	*38.80%*	*38.77%*

presented. As with the SDS-PAGE analysis, relative protein content is presented, to prevent the results being compounded with differences in protein content, and to remove some of the inaccuracies possible in this form of protein quantification.

Grain gluten content was calculated by subtracting the summed measurements for the F1–F4 protein fractions from the total protein detected (F1–F5), and was found to be significantly higher under both the elevated post-anthesis temperature ($F_{1,4}$=32.73, P=0.005, SED=0.4385 on 4 DF) (figure 6.7a) and the high nitrogen input treatments ($F_{1,4}$=137.15, P<0.001, SED=0.0934 on 4 DF) (figure 6.7b), with no significant interaction between these factors ($F_{1,4}$=2.08, P=0.222). Of the two factors, temperature had the greatest effect on grain gluten content, with approximately double the difference between the two temperature treatments when compared to the nitrogen treatments. The gluten content data is presented in full in table 6.5.

Reliable estimations of bread-making quality can be made by analysing the ratios between

Table 6.5: Full dataset for grain gluten content at maturity from the controlled-environment experiment, presented per treatment combination, averaged over experimental blocks. Expressed as the total of SE-HPLC fractions F1–F4 as a percentage of the total protein detected by SE-HPLC (F1–F5).

Treatment	Gluten content (%)
20°C, low-nitrogen	82.31
20°C, high-nitrogen	83.53
28°C, low-nitrogen	84.95
28°C, high-nitrogen	85.91

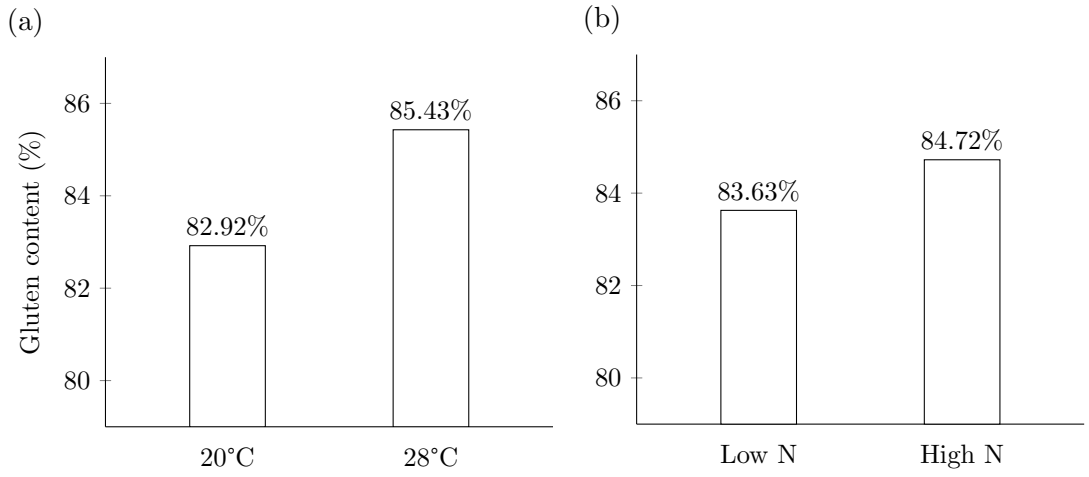


Figure 6.7: **Gluten content of mature grain is increased by high nitrogen and high temperatures.** Proportion of protein detected by SE-HPLC as gluten protein (F1–F4 as percentage of F1–F5), grouped by the single significant effects of (a) temperature treatment, LSD of 0.12; and (b) nitrogen treatment, LSD (at the 5% level) of 0.26.

.

gluten protein components (Millar 2003). The two ratios of interest are the F1/F2 ratio, which represents the ratio between large HMW- and LMW-glutenins polymers, and the (F3+F4)/F1 ratio, which is the ratio between small gliadin polymers and HMW-glutenins. Higher F1/F2, and lower (F3+F4)/F1 numbers are associated with superior bread-making quality. All ratios were calculated using the relative protein content values (% of total protein).

The ratio between HMW- and LMW-glutenins (F1/F2) was determined by an interaction between post-anthesis temperature and nitrogen treatments ($F_{1,4}$=10.01, P=0.034, SED=0.1387 on 7.46DF): both elevated post-anthesis temperature and high nitrogen input resulted in an increase in the F1/F2 ratio, but the effect of nitrogen input was only significant under control temperatures (figure 6.8a). A two-way interaction between temperature and nitrogen was also found for the ratio between gliadins and HMW-glutenins ((F3+F4)/F1) ($F_{1,4}$=8.45, P=0.044, SED=0.2360 on 7.13DF), with both elevated temperature and high nitrogen input *decreasing* the (F3+F4)/F1 ratio, with the effect of nitrogen again only significant at the 5% level under the control temperature treatment (figure 6.8).

Figure 6.9 shows a plot of the F1/F2 ratio against the (F3+F4)/F1 ratio, and provides a visual representation of how predicted bread-making quality differs between grain from the different treatments. Datapoints to the lower right of the graph signify superior bread-making

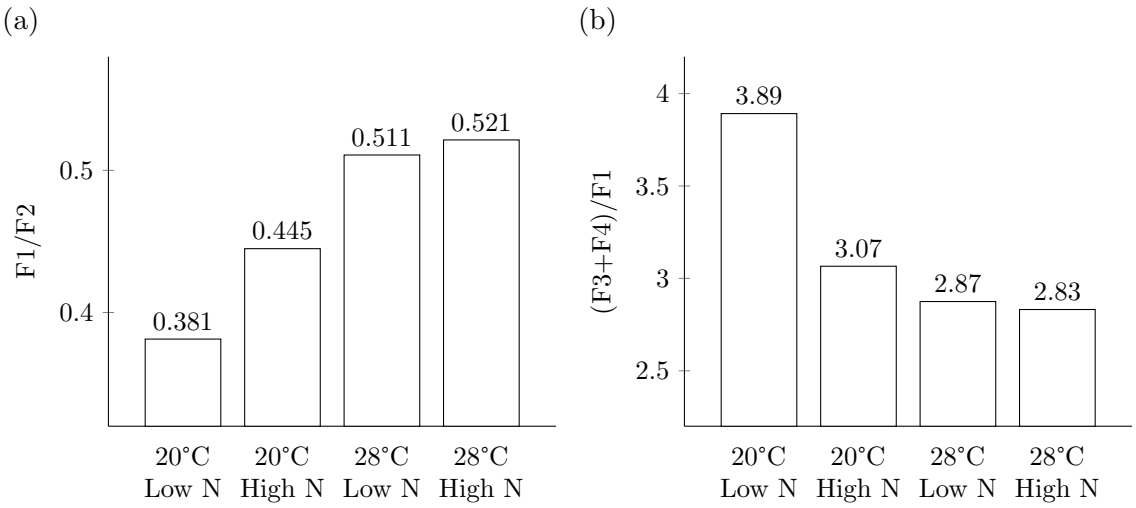


Figure 6.8: **Elevated post-anthesis temperature and increased nitrogen input increase the ratio between HMW-glutenins, and decrease the ratio between gliadins and HMW-glutenins.** The combined effect of temperature and nitrogen on the ratios between the relative amount of SE-HPLC protein fractions related to bread-marking quality from the controlled-environment experiment: (a) HMW- to LMW-glutenin ratio (F1/F2), LSD (at the 5% level) of 0.033 for comparing within the same temperature treatment, and 0.032 for all other comparisons; and (b) gliadin to HMW-glutenin ratio ((F3+F4)/F1), LSD of 0.53 for comparing within the same temperature treatment, and 0.56 for all other comparisons.

.

quality, and the figure shows that both increased temperature and higher nitrogen input improves the predicted bread-making quality of the grain. Furthermore, the effect of the two-way interactions between temperature and nitrogen treatments can also be observed, with a greater separation of nitrogen treatments under the control temperature treatment.

Analysis of the individual protein fractions was performed on the relative data (figure 6.10), which expressed the fractions as a percentage of the total protein detected. As with the SDS-PAGE data, this allows differences in protein composition to be identified, without having the results affected by the differences in total protein content that were also observed. The results

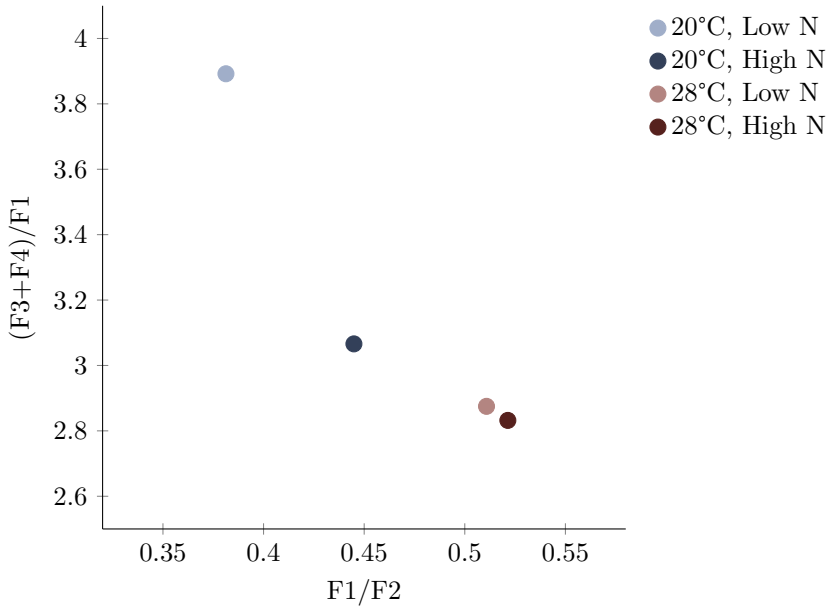


Figure 6.9: **Predicted bread-making quality is improved by high nitrogen input and elevated post-anthesis temperature.** Plot of F1/F2 against (F3+F4)/F1 SE-HPLC protein fractions for each combination of temperature and nitrogen treatment used in the controlled-environment experiment. F1/F2 represents the ratio between HMW- and LMW-glutenins, and (F3+F4)/F1 represents the ratio between gliadins and HMW-glutenins. Points to the lower right of the graph indicate superior bread-making quality. All ratios were calculated with the relative protein content data (as a % of the total protein).

.

Table 6.6: Results from the ANOVA analyses of SE-HPLC data from the environment experiment. Table shows the significant effects of all treatment combinations on individual protein fractions from the SE-HPLC analysis. Effects to interpret are in shown in bold.

Fixed term	F1	F2	F3	F4	F5
Temperature	0.004	0.136	**0.029**	0.368	**0.005**
Nitrogen	0.010	0.091	**0.027**	0.136	**<0.001**
Temperature.Nitrogen	**0.031**	0.182	0.058	0.056	0.222

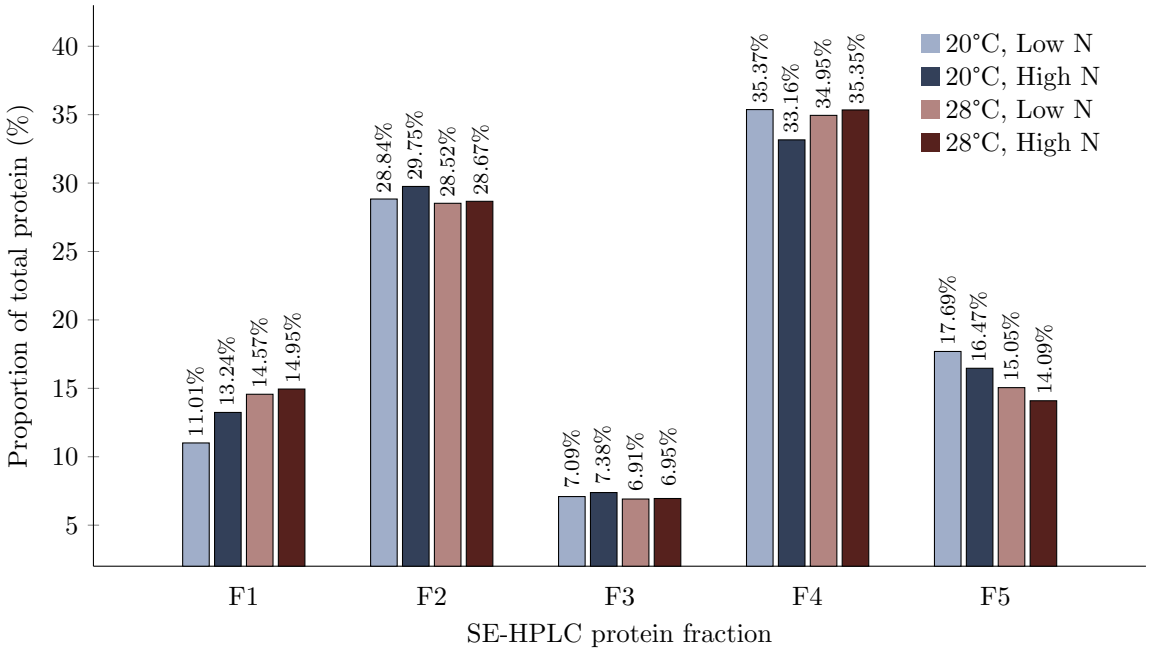


Figure 6.10: **The relative abundance of SE-HPLC protein fractions are differentially affected by nitrogen supply and post-anthesis temperature.** Relative protein content of mature grain by SE-HPLC, grouped by each combination of temperature and nitrogen treatment from the controlled environment experiment. Protein fractions represent HMW-glutenins (F1), LMW-glutenins (F2), omega-gliadins (F3), alpha-, beta-, and gamma-gliadins (F4), and albumins and globulins (F5).

.

of the ANOVA analyses of the protein composition data is presented in table 6.6.

The proportion of HMW-glutenins in the mature grain was increased by both elevated post-anthesis temperature, and increased nitrogen input, with an interaction between these two factors ($F_{1,4}$=10.70, P<0.031, SED=0.5162 on 4 DF) whereby nitrogen input had a greater effect under control temperatures than under elevated temperatures. There was no evidence of an effect of temperature on the concentration of LMW-glutenins ($F_{1,4}$=3.47, P=0.136), but there was some evidence that high nitrogen input increases the concentration of LMW-glutenins ($F_{1,4}$=4.91, P=0.091) although not significant at the 5% level. The concentration of omega-gliadins at maturity was increased by high nitrogen input ($F_{1,4}$=11.65, P=0.027, SED=0.0487 on 4 DF), but decreased by elevated post-anthesis temperature ($F_{1,4}$=12.05, P=0.029, SED=0.0925 on 4 DF), with some evidence that the effect of nitrogen was greater under control temperatures ($F_{1,4}$=6.90, P=0.058). For alpha-, beta-, and gamma-gliadins, there was no significant effect of temperature ($F_{1,4}$=1.03, P=0.368) or nitrogen ($F_{1,4}$=3.47, P=0.136), but there was some evidence of an interaction between these two factors ($F_{1,4}$=7.15, P=0.056). This interaction is again identifying an increased response to nitrogen under control temperature conditions, whereby the proportion of alpha-, beta-, and gamma-gliadins is increased by high nitrogen under control temperatures, but not under elevated temperatures.

Finally, as the inverse of the gluten content results, the proportion of albumin and globulins in the mature grain are decreased by both elevated temperature ($F_{1,4}$=32.73, P=0.005, SED=0.4385 on 4 DF) and high nitrogen input ($F_{1,4}$=137.15, P<0.001, SED=0.0934 on 4 DF).

6.3. **WGIN Diversity Field Experiment**

To investigate the effect of nitrogen application and year-to-year climatic variation on the protein composition of different wheat genotypes, grain from the WGIN diversity field experiment was sampled and analysed. Samples were taken over three years with varying weather conditions during grain-filling (discussed in chapter 3), from four commercial wheat genotypes grown under either low (100kg-N/ha) or abundant (350kg-N/ha) nitrogen levels. Analysis of the protein composition of grain samples collected as part of the WGIN diversity field experiment was similar to the analysis completed for the controlled-environment experiment: nitrogen content analysis was used to measure protein concentration, with protein composition determined using both SDS-PAGE and SE-HPLC. However, whilst nitrogen content was measured both during grain-filling and at maturity, SDS-PAGE and SE-HPLC analysis was only completed on mature grain. Data for the calculated grain protein content is not presented, since this data was compounded with the TGW data (from which it is calculated), making the identification of genuine effects impossible.

6.3.1. Protein concentration

The analysis of grain protein concentration (nitrogen content multiplied by 5.7) during grain-filling identified two three-way interactions: interactions between genotype, nitrogen treatment, and timepoint of measurement ($F_{15,213}$=4.21, P<0.001, SED=0.2535 on 251.6 DF), and between year of experiment, nitrogen treatment, and timepoint ($F_{10,213}$=4.93, P<0.001, SED=0.2418 on 74.24 DF) were found. These two interactions describe the effects of all treatments in the experiment, and the complete dataset is presented in figure 6.11. The first interaction, between genotype, nitrogen, and timepoint, shows that different genotypes accumulate nitrogen during grain-filling differently under each of the two nitrogen treatments, and that this response was comparable between years. Analysis of the LSD values from the second interaction (between year, nitrogen, and timepoint) show that this interaction is describing the response to the high nitrogen treatment in 2017 (the hottest year), where grain protein concentration is increased during mid to late grain-filling.

For protein concentration measured at maturity, there was a significant interaction between year of experiment and genotype ($F_{5,31}$=10.82, P<0.001, SED=0.2906 on 21.47 DF) (figure 6.12a), which primarily describes the differential response shown by different genotypes in 2016: Cadenza showed an increase in protein concentration from 2015 to 2016 to 2017, whilst Hereward showed a marked decrease between 2015 and 2016, before increasing again in 2017, and Istabraq and Soissons achieved similar protein concentrations in 2015 and 2016, before increasing in 2017. Overall, the lowest protein concentration was recorded in 2016,

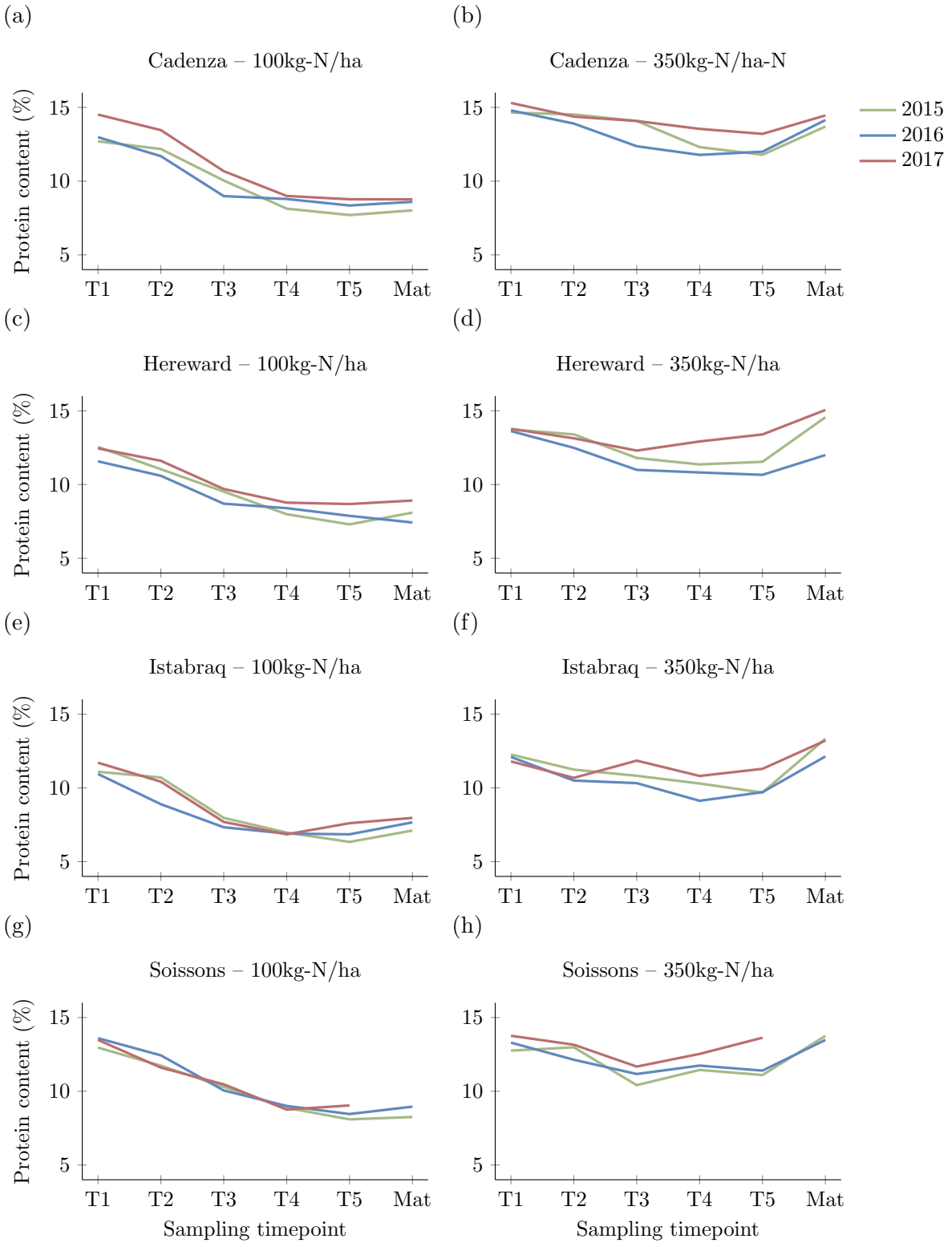


Figure 6.11: **Grain protein concentration during grain-filling is increased by nitrogen input differentially by different genotypes, and over different years.** Grain protein concentration through development for Cadenza, Hereward, Istabraq, and Soissons under low (100kg-N/ha) and high (350kg-N/ha) nitrogen input in the 2015 (green), 2016 (blue), and 2017 (red) WGIN diversity field experiment. Calculated from nitrogen content data using a conversion factor of 5.7, and expressed as percentage dry matter.

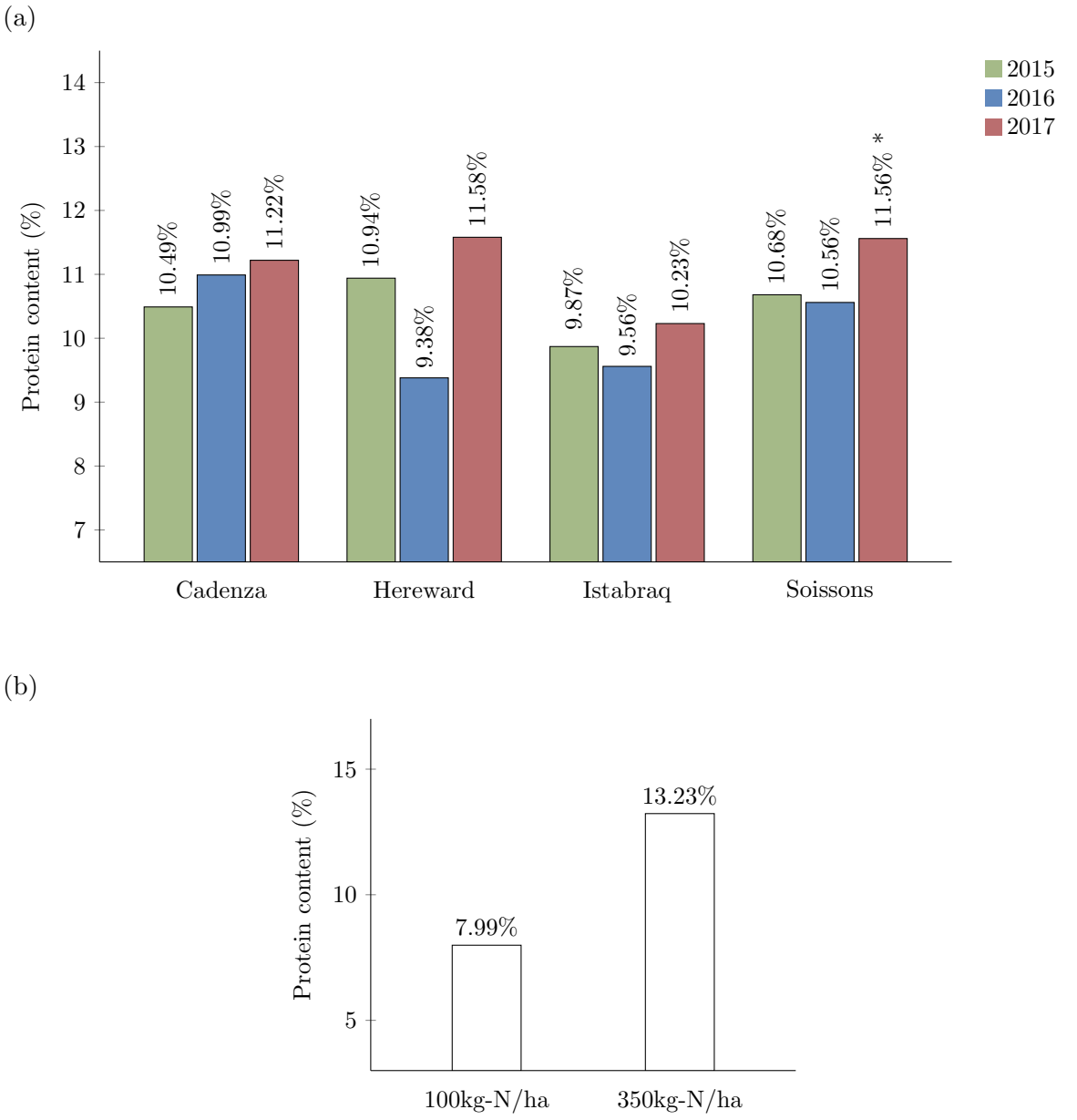


Figure 6.12: **Mature grain protein concentration is increased in hotter years, and by high nitrogen application rate.** Protein concentration data from the WGIN diversity field experiment, grouped by (a) genotype and year combination, LSD (at the 5% level) of 0.50 for comparing within the same year, and 0.60 for all other comparisons; and (b) nitrogen treatment, LSD of 0.44. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

Table 6.7: Full dataset for grain protein concentration at maturity from the WGIN diversity field trial experiment, averaged over experimental blocks. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		Protein content (%)		
Nitrogen input	Genotype	2015	2016	2017
100kg-N/ha	Cadenza	7.75	8.31	8.47
	Hereward	7.82	7.17	8.61
	Istabraq	6.87	7.40	7.70
	Soissons	8.06	8.65	9.06*
350kg-N/ha	Cadenza	13.24	13.67	13.97
	Hereward	14.07	11.59	14.54
	Istabraq	12.88	11.72	12.76
	Soissons	13.29	13.02	14.05*

and the highest in 2017. Additionally, a single effect from nitrogen treatment was identified ($F_{1,6}$=865.23, P<0.001, SED=0.1783 on 6 DF) (figure 6.12b), which simply describes an increase in grain protein concentration under the high (350kg-N/ha) nitrogen treatment. The full dataset for protein concentration at maturity is presented in table 6.7.

6.3.2. SDS-PAGE

As part of the WGIN diversity field experiment, SDS-PAGE was used to measure protein composition at maturity only. Since these measurements were taken from multiple genotypes, each of which produce different gluten protein subunits, proteins were grouped rather than analysed individually. As with SE-HPLC, proteins were categorised as either alpha-, beta-, and gamma-gliadins, HMW-glutenins, LMW-glutenins, or omega-gliadins. As with the controlled-environment experiment, only the relative levels of each protein group within a sample were analysed.

Preliminary CV analysis of the SDS-PAGE data showed a strong separation of the two nitrogen treatments, but failed to show a clear separation between the different genotypes or years (figure 6.13). This suggests that nitrogen input level had the greatest effect on protein composition, and that genotype and year of harvest have weaker influences. In the CV analysis presented in figure 6.13, CV1 accounts for 86.81% of variation, and is primarily separating the two nitrogen treatments. Whilst both CV1 and CV2 (which represents 7.94% of variation) show some separation between genotypes and years of experiment. Of the four genotypes analysed, Soissons shows the least separation, whilst Hereward shows the most, which could indicate that Soissons may be less susceptible to changes in nitrogen input, with regards to protein composition, than Hereward. Vector loadings for each protein group are presented in figure 6.14, and show the influence each protein group had on the separation of datapoints shown in the CV analysis. From these vector loadings, omega-gliadins are the protein group most affected by changes to nitrogen input, closely followed by LMW- and HMW-glutenins, with alpha-, beta-, and gamma-gliadins the least affected.

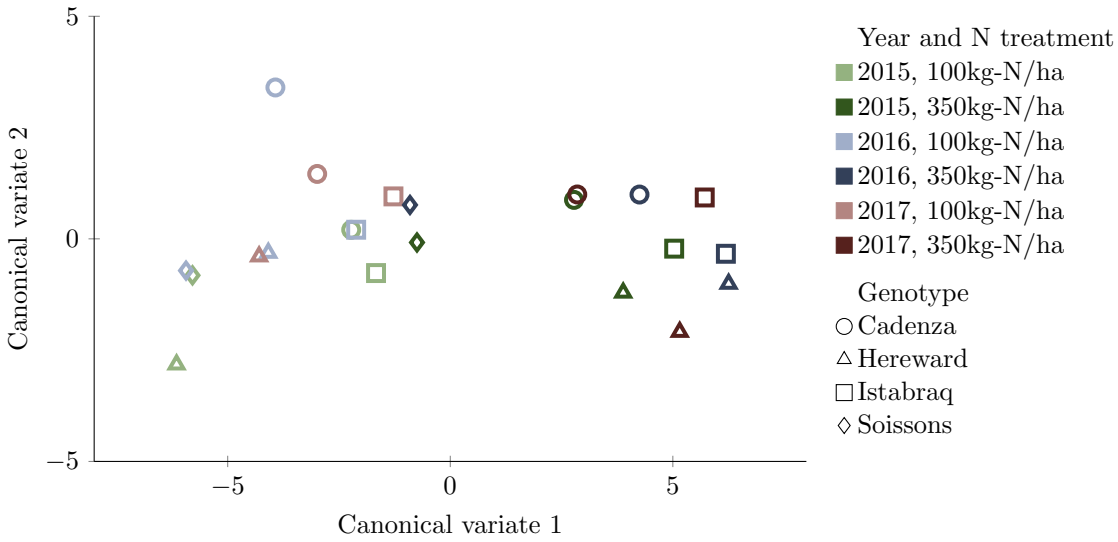


Figure 6.13: **Grain protein composition is differentially affected by nitrogen application rate.** CV analysis of SDS-PAGE data from the WGIN diversity field experiment. Analysis separates the effects of year of experiment, genotype, and nitrogen treatment based on the overall protein composition. CV1 primarily represents the effect of nitrogen treatment whilst the effect of genotype and year of experiment are not clearly represented by either CV1 or CV2. CV1 and CV2 account for 86.81% and 7.94% of the observed variation respectively.

.

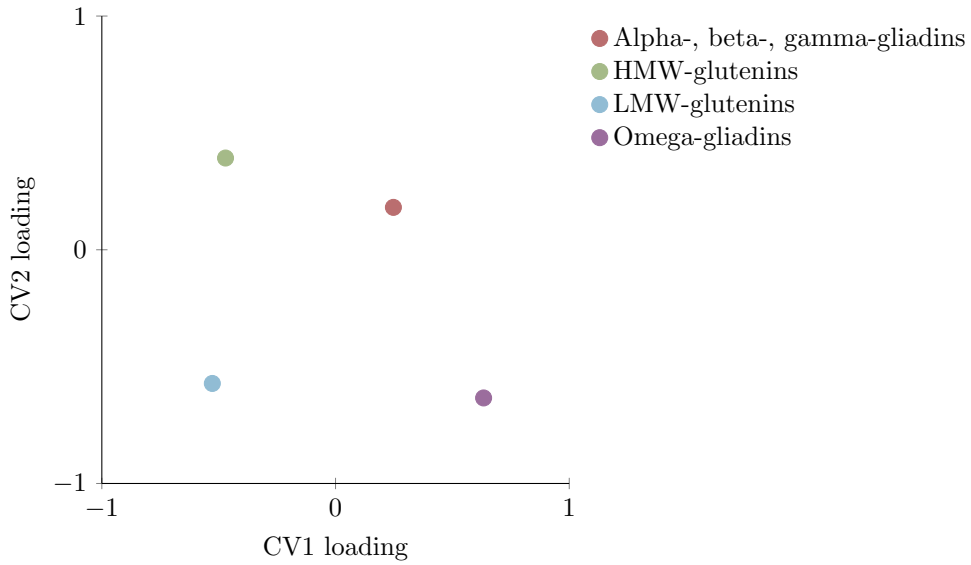


Figure 6.14: **Omega-gliadins and LMW-glutenins show strong influence on the separation of nitrogen treatments by CV analysis.** Vector loadings for CV analysis of SDS-PAGE data from the WGIN diversity field experiment experiment. Loadings indicate the influence each protein group has on the CV analysis presented in figure 6.13. Omega-gliadins, LMW-glutenins, and HMW-glutenins have the strongest influence on CV1, which represents the separation of nitrogen treatments.

.

Table 6.8: Results from the REML analyses of SDS-PAGE data from the WGIN diversity field experiment. Table shows the significant effects of all treatment combinations on each protein group. Effects to interpret are in shown in bold. The significant three- and two-way interactions for alpha-, beta-, and gamma-gliadins and HMW-glutenins groups were investigated, and found to be unsuitable interpretations of the data.

Fixed term	Alpha-, beta-, gamma-gliadins	HMW-glutenins	LMW-glutenins	Omega-gliadins
Year	0.380	0.309	0.271	0.353
Genotype	**<0.001**	**<0.001**	<0.001	<0.001
Nitrogen	0.952	0.799	<0.001	<0.001
Year.Genotype	0.111	0.830	0.201	0.198
Year.Nitrogen	0.267	0.037	0.177	0.722
Genotype.Nitrogen	0.309	0.472	**<0.001**	**<0.001**
Year.Genotype.Nitrogen	0.023	0.025	0.610	0.200

Separate REML analyses was used to determine the effect that genotype, nitrogen, and year of experiment had on the relative proportion of each protein group measured by SDS-PAGE. The results of these analyses are presented in table 6.8. For the alpha-, beta-, and gamma-gliadins and HMW-glutenins protein groups, borderline significant interactions were identified between year of experiment, genotype, and nitrogen treatment. However upon inspection of the means with the LSD values, these interactions fail to identify significant differences between treatment combinations of interest. Likewise for the HMW-glutenins a two-way interaction between year and nitrogen was identified, but again comparison of the means revealed no significant differences. Therefore for these two protein groups, the single effect of genotype is interpreted instead.

The proportion of alpha-, beta-, and gamma- gliadins showed a strong effect from genotype ($F=82.59$, $P<0.001$, SED=0.3927 on 31.8 DF) (figure 6.15a), and was highest in Istabraq, followed by Cadenza, Hereward and Soissons, with the difference between each genotype significant at the 5% level.

There was also a strong effect of genotype on the proportion of detected HMW-glutenins ($F=39.04$, $P<0.001$, SED=0.4003 on 31.8 DF) (figure 6.15b), with the highest proportion detected in Soissons, followed by Cadenza, Hereward and Istabraq.

For the LMW-glutenins, a two-way interaction between genotype and nitrogen treatment was identified ($F=9.41$, $P<0.001$, SED=0.6188 on 32.2 DF) (figure 6.15c), with no effect of year ($F=1.46$, $P=0.271$). Increased nitrogen input decreased the proportion of LMW-glutenins in all genotypes, but the magnitude of this effect differed between the different genotypes: Hereward showed the greatest response, followed by Istabraq, whilst Cadenza and Soissons showed much smaller responses.

A two-way interaction between genotype and nitrogen was also found for the concentration of omega-gliadins present ($F=21.01$, $P<0.001$, SED=0.5494 on 31.9 DF) (figure 6.15c), again

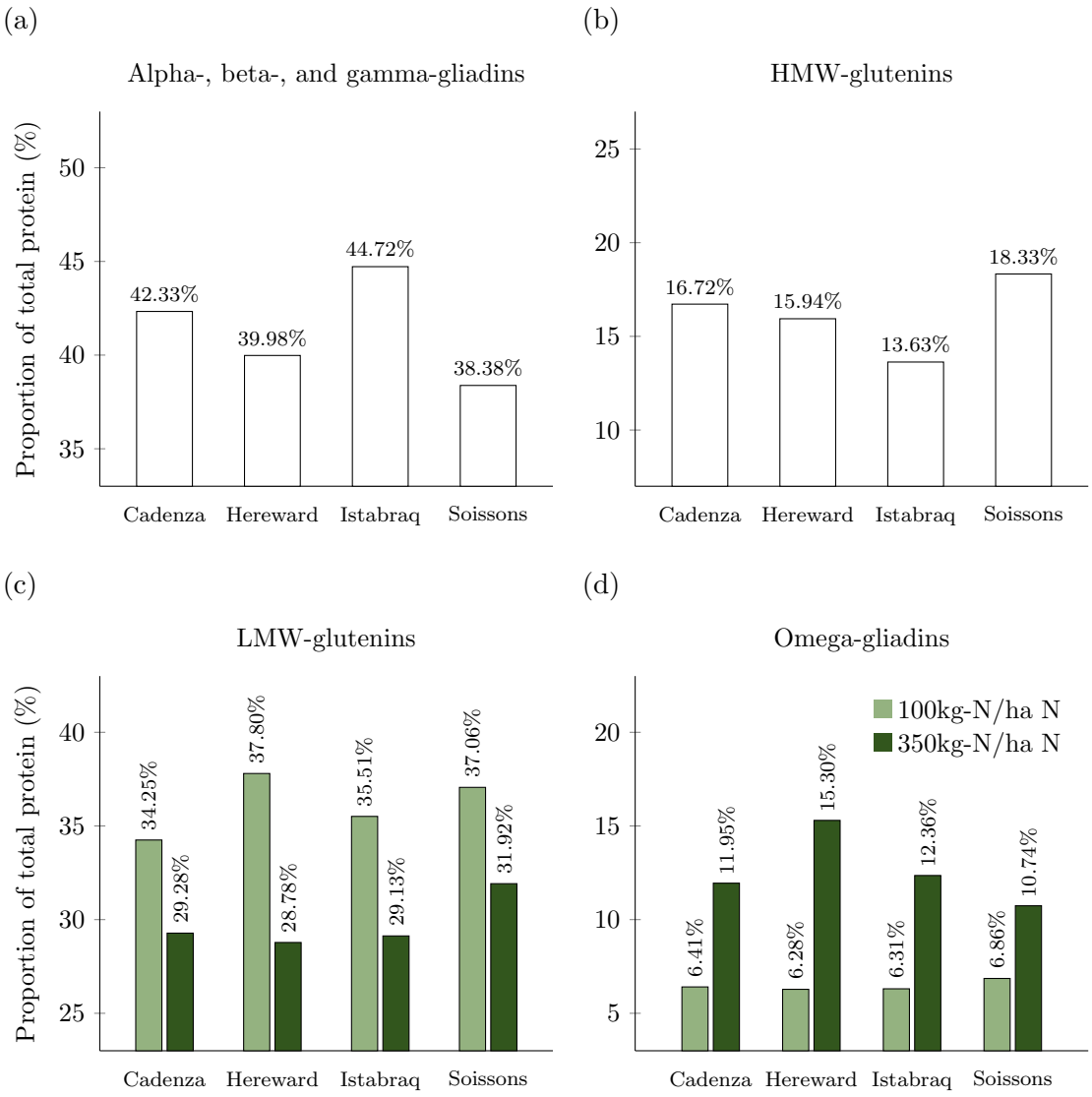


Figure 6.15: **Nitrogen input interacts with genotype to determine grain protein composition.** Proportion of protein groups as percentage of the total protein detected by SDS-PAGE from the WGIN diversity field experiment: (a) alpha-, beta-, and gamma-gliadins by genotype, LSD (at the 5% level) of 0.80; (b) HMW-glutenins by genotype, LSD of 0.82; (c) LMW-glutenins by combination of genotype and nitrogen treatment, LSD of 1.38 for comparing within the same nitrogen treatment, and 1.41 for all other comparisons; (d) omega-gliadins by genotype and nitrogen treatment, LSD of 1.01 for comparing within the same nitrogen treatment, and 1.40 for all other comparisons.

.

Table 6.9: Full SDS-PAGE protein composition dataset from the WGIN diversity field trial experiment. Individual proteins detected were assigned to one of four groups, and the total of each protein group as a percentage of the total detected protein is presented.

Year	Nitrogen input	Genotype	Alpha-, beta-, gamma-gliadins	HMW-glutenins	LMW-glutenins	Omega-gliadins
2015	100kg-N/ha	Cadenza	46.76%	11.74%	37.05%	4.46%
		Hereward	42.74%	11.69%	41.81%	3.76%
		Istabraq	47.70%	9.99%	37.92%	4.39%
		Soissons	41.51%	14.79%	38.96%	4.74%
	350kg-N/ha	Cadenza	44.48%	14.71%	30.70%	10.12%
		Hereward	42.39%	14.34%	30.79%	12.48%
		Istabraq	47.24%	11.43%	30.96%	10.37%
		Soissons	41.73%	15.79%	33.62%	8.86%
2016	100kg-N/ha	Cadenza	38.36%	22.02%	31.55%	8.07%
		Hereward	36.45%	19.74%	34.49%	9.32%
		Istabraq	41.01%	16.64%	34.22%	8.13%
		Soissons	33.90%	21.32%	35.42%	9.35%
	350kg-N/ha	Cadenza	41.06%	17.55%	27.84%	13.55%
		Hereward	37.61%	18.42%	26.44%	17.54%
		Istabraq	41.14%	16.22%	27.41%	15.23%
		Soissons	35.98%	21.20%	30.50%	12.32%
2017	100kg-N/ha	Cadenza	42.02%	17.12%	34.17%	6.70%
		Hereward	41.87%	15.26%	37.12%	5.76%
		Istabraq	44.86%	14.36%	34.38%	6.41%
		Soissons	—	—	—	—
	350kg-N/ha	Cadenza	41.32%	17.20%	29.31%	12.17%
		Hereward	38.82%	16.19%	29.11%	15.87%
		Istabraq	46.37%	13.14%	29.02%	11.46%
		Soissons	—	—	—	—

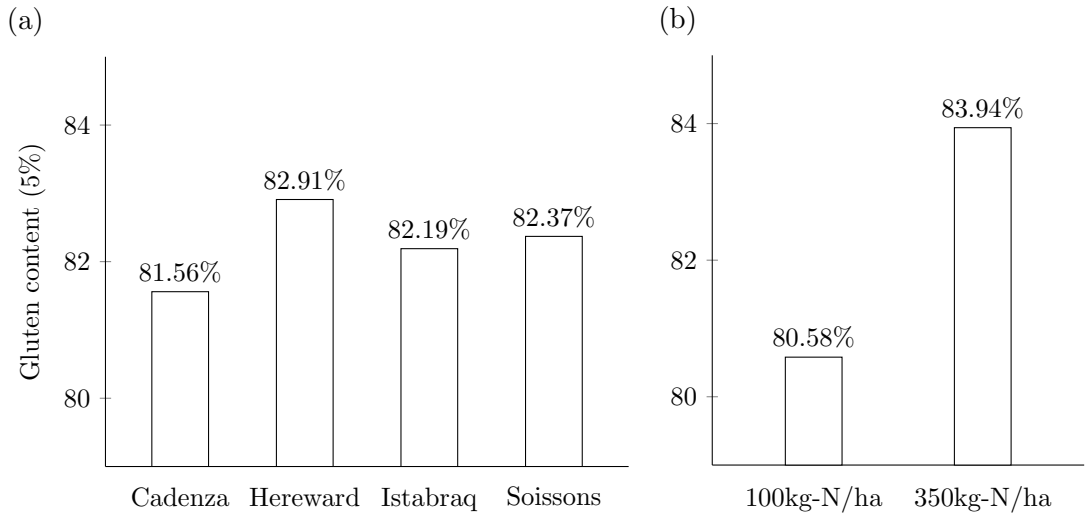


Figure 6.16: **Gluten content varies between genotypes, and is increased by high nitrogen input.** Gluten content as a percentage of the total protein detected by SE-HPLC at maturity from the WGIN diversity field experiment. (a) mean gluten content for each genotype, LSD (at the 5% level) of 0.75; (b) gluten content for each nitrogen treatment, LSD of 2.83.

.

with no effect of year (F=1.28, P=0.353). The proportion of omega-gliadins was significantly increased by the high nitrogen treatment in all genotypes, and as with LMW-glutenins, Hereward showed the strongest response to nitrogen, followed by Istabraq, Cadenza, and Soissons. The full dataset is presented in table 6.9.

6.3.3. SE-HPLC

SE-HPLC analysis was completed on mature grain from the WGIN diversity field experiment using the same methodology as used in the controlled-environment experiment, with

Table 6.10: Full gluten content dataset from the WGIN diversity field trial experiment, presented per treatment combination, averaged over experimental blocks. Expressed as the total of SE-HPLC fractions F1–F4 as a percentage of the total protein detected by SE-HPLC (F1–F5). Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		Gluten content (%)		
Nitrogen input	Genotype	2015	2016	2017
100kg-N/ha	Cadenza	80.39	82.12	76.67
	Hereward	82.00	81.37	79.14
	Istabraq	81.67	82.55	78.47
	Soissons	80.62	83.46	78.45*
350kg-N/ha	Cadenza	84.89	83.82	81.48
	Hereward	87.24	84.88	82.85
	Istabraq	85.66	84.13	80.66
	Soissons	85.65	84.46	81.62*

gluten content, F1/F2 and (F3+F4)/F1 ratios, and individual protein fraction abundance presented.

Relative gluten content was significantly different between genotypes ($F_{3,32}$=4.56, P=0.009, SED=0.369 on 32 DF) (figure 6.16a), with Hereward having the highest gluten content and Cadenza the lowest. Gluten content was significantly affected by nitrogen input ($F_{1,6}$=8.46, P=0.027, SED=1.158 on 6 DF) (figure 6.16b), with the 350kg-N/ha nitrogen treatment increasing gluten content from 80.58% to 83.94% when compared to the 100kg-N/ha treatment. The full dataset is presented in table 6.10.

The ratio between large HMW- and LMW-glutenin polymers (F1/F2) was not found to be significantly affected by year ($F_{2,6}$=1.95, P=0.223), nitrogen ($F_{1,6}$=1.57, P=0.257), or genotype ($F_{3,32}$=0.03, P=0.993), and no combination of these factors had a significant effect. Likewise, for the ratio between small gliadin polymers and large HMW-glutenin polymers ((F3+F4)/F1), no significant effect was found for year ($F_{2,6}$=2.31, P=0.181), nitrogen ($F_{1,6}$=0.10, P=0.768), or genotype ($F_{3,32}$=0.76, P=0.526), nor for any combination of these factors. The full dataset of F1/F2 and (F3+F4)/F1 ratios is presented in table 6.11.

Analysis of the individual protein fractions detected by SE-HPLC was by ANOVA, and the results of these analyses are presented in table 6.12, whilst the full dataset is presented in table 6.13.

No significant effect on the proportion of HMW-glutenins (F1) in the mature grain was found from year ($F_{2,6}$=2.91, P=0.131), nitrogen ($F_{1,6}$=1.11, P=0.333), genotype ($F_{3,32}$=0.16, P=0.922). Likewise, analysis of the results for the LMW-glutenin (F2) fraction showed similar results, with no significant effect from year ($F_{2,6}$=1.45, P=0.307), nitrogen ($F_{1,6}$=2.61, P=0.157), or genotype ($F_{3,32}$=1.77, P=0.173).

The proportion of detected omega-gliadins (F3) was affected by genotype ($F_{3,32}$=6.46, P=0.002,

Table 6.11: Full dataset of ratios between SE-HPLC protein fractions F1/F2 and (F3+F4)/F1 from the WGIN diversity field trial experiment. Lower F1/F2 and higher (F3+F4)/F1 values are associated with superior bread-making quality. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

		F1/F2 ratio			(F3+F4)/F1 ratio		
Nitrogen input	Genotype	2015	2016	2017	2015	2016	2017
100kg-N/ha	Cadenza	0.86	0.78	0.63	2.35	2.89	4.16
	Hereward	0.91	0.86	0.55	1.97	3.77	4.38
	Istabraq	0.89	0.79	0.66	2.20	2.73	4.07
	Soissons	0.68	1.01	0.61*	2.72	2.18	4.00*
350kg-N/ha	Cadenza	0.92	0.80	0.89	2.33	2.87	4.07
	Hereward	1.01	0.97	0.58	2.19	2.93	4.15
	Istabraq	0.91	0.84	0.81	2.49	2.72	4.08
	Soissons	0.88	0.91	0.73*	2.49	2.36	3.89*

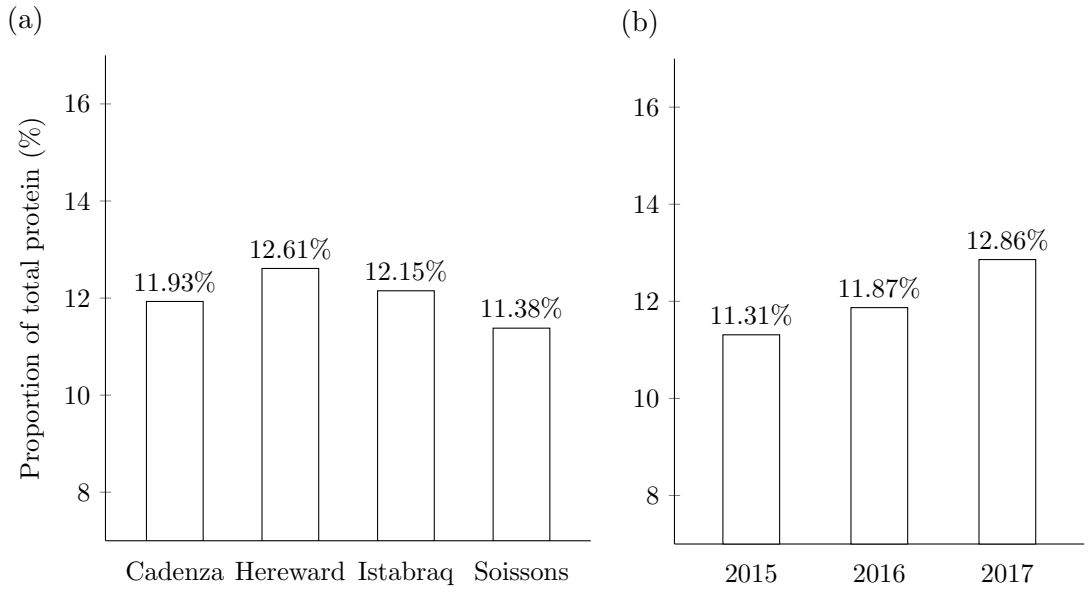


Figure 6.17: **Omega-gliadin content of mature grain varies between genotypes, and was highest in the hottest year.** Relative omega-gliadin (F3) content of mature grain by SE-HPLC from the WGIN diversity field experiment grouped by (a) genotype, LSD (at the 5% level) of 0.58; and (b) year, LSD of 1.13.

.

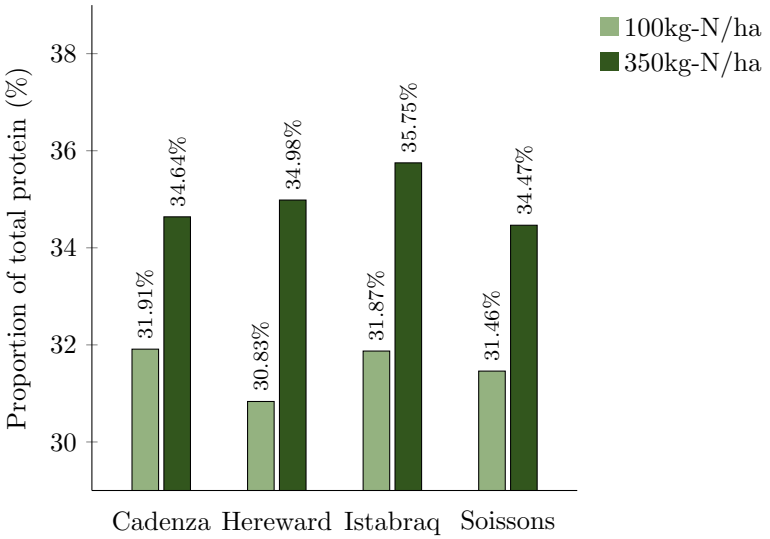


Figure 6.18: **Alpha-, beta-, and gamma-gliadin content of the mature grain is differentially increased by high nitrogen input in different genotypes.** Relative alpha-, beta-, and gamma-gliadin (F4) content at maturity from the WGIN diversity field experiment, averaged across years to show the significant interaction between genotype and nitrogen treatment, LSD of 0.78 for comparing within the same nitrogen treatment, and 2.61 for all other comparisons.

.

Table 6.12: Results from the ANOVA analyses of SE-HPLC data from the WGIN diversity field experiment. Table shows the significant effects of all treatment combinations on each protein group. Effects to interpret are in shown in bold. The significant two-way interaction between year and genotype for protein group F4 was investigated, and found to be an unsuitable interpretation of the data.

Fixed term	F1	F2	F3	F4	F5
Year	0.131	0.307	**0.040**	0.215	0.082
Genotype	0.922	0.173	**0.002**	0.008	**0.009**
Nitrogen	0.333	0.157	0.084	0.018	**0.027**
Year.Genotype	0.129	0.121	0.806	0.025	0.124
Year.Nitrogen	0.960	0.163	0.799	0.235	0.646
Genotype.Nitrogen	0.948	0.795	0.528	**0.038**	0.180
Year.Genotype.Nitrogen	0.750	0.949	0.270	0.455	0.467

SED=0.285 on 32 DF) (figure 6.17a), and year ($F_{2,6}$=5.77, P=0.040, SED=0.463 on 6 DF) (figure 6.17b), with some evidence of an effect from nitrogen ($F_{1,6}$=4.28, P=0.084). Of the three years, the highest proportion of omega-gliadins was found in 2017, and the lowest in 2015, with the difference between these two years the only significant difference. Of the four genotypes, Hereward had the highest proportion of omega-gliadins, followed by Istabraq and Cadenza, with Soissons having the lowest, and whilst not significant at the 5% level, the proportion of omega-gliadins was slightly reduced by the high nitrogen treatment when averaging over genotypes and years of experiment.

With regards to alpha-, beta-, and gamma-gliadins, two significant two way interactions were found, one between genotype and year ($F_{5,32}$=2.99, P=0.025, SED=1.3920 on 7.16 DF), and another between genotype and nitrogen treatment ($F_{3,32}$=3.17, P=0.038, SED=1.1098 on 7.22 DF). However, upon further inspection the interaction between genotype and year was found to be a poor fit for the data as the differences between years were not significant at the 5% level, and since variation between genotypes is accounted for by the interaction between genotype and nitrogen treatment, this interaction was dropped. Looking at the results of the genotype by nitrogen interaction (figure 6.18), the 350kg-N/ha nitrogen treatment significantly increased the proportion of alpha-, beta-, and gamma-gliadins for every genotype, with Hereward and Istabraq showing greater responses than Cadenza and Soissons.

The effect of genotype, year, and nitrogen treatment on the proportion of albumin and globulins (F5) mirrors the results for the gluten content analysis, and are therefore not presented.

6.4. Discussion

6.4.1. Controlled-environment experiment

The nitrogen content data collected as part of the controlled-environment experiment allows for the total protein concentration of the grain to be tracked throughout grain-filling, and also provides a final protein content for the mature grain that can be used to compare the

Table 6.13: Full SE-HPLC protein composition dataset from the WGIN diversity field trial experiment. Relative concentration of HMW-glutenins (F1), LMW-glutenins (F2), omega-gliadins (F3), alpha-, beta-, and gamma-gliadins (F4), and albumins and globulins (F4) as a percentage of the total protein (F1–F5) are presented. Predicted values from the ANOVA model are presented for Soissons in 2017 (marked *).

Year	Nitrogen input	Genotype	F1	F2	F3	F4	F5
2015	100kg-N/ha	Cadenza	18.13%	21.02%	11.73%	29.51%	19.61%
		Hereward	20.29%	22.60%	11.69%	27.42%	18.00%
		Istabraq	18.94%	21.42%	12.04%	29.27%	18.33%
		Soissons	16.11%	24.16%	10.94%	29.41%	19.38%
	350kg-N/ha	Cadenza	19.22%	20.97%	10.64%	34.07%	15.11%
		Hereward	20.87%	20.62%	11.58%	34.16%	12.76%
		Istabraq	19.04%	19.72%	11.03%	35.86%	14.34%
		Soissons	18.89%	21.57%	10.82%	34.38%	14.35%
2016	100kg-N/ha	Cadenza	15.89%	20.41%	12.27%	33.55%	17.88%
		Hereward	15.17%	19.25%	13.35%	33.59%	18.63%
		Istabraq	16.68%	21.06%	11.62%	33.19%	17.45%
		Soissons	19.99%	19.96%	10.90%	32.60%	16.54%
	350kg-N/ha	Cadenza	16.92%	21.39%	11.26%	34.25%	16.18%
		Hereward	17.94%	20.48%	12.20%	34.26%	15.12%
		Istabraq	17.17%	20.51%	12.24%	34.20%	15.87%
		Soissons	19.08%	20.99%	11.15%	33.24%	15.54%
2017	100kg-N/ha	Cadenza	11.90%	19.09%	13.01%	32.67%	23.33%
		Hereward	11.95%	21.83%	13.86%	31.50%	20.86%
		Istabraq	12.51%	19.02%	13.78%	33.16%	21.53%
		Soissons	12.66%*	21.08%*	12.35%*	32.36%*	21.55%*
	350kg-N/ha	Cadenza	15.17%	18.06%	12.66%	35.50%	18.52%
		Hereward	12.20%	21.17%	12.96%	36.52%	17.15%
		Istabraq	13.96%	17.32%	12.19%	37.19%	19.34%
		Soissons	14.23%*	19.51%*	12.09%*	35.78%*	18.38%*

grain to that grown in the field. Furthermore, by comparing protein concentration with protein content, we can determine whether the changes in grain protein concentration are due to a change in the accumulation of protein, or other cellular components such as starch. Elevated temperature during grain-filling increased the *concentration* of protein in the grain throughout development, and this trend continued to maturity. Increasing nitrogen input had a similar effect, resulting in an increase in grain protein concentration both throughout development, and at maturity. Additionally, the effect of nitrogen input was greater under the high-temperature treatment, producing a greater increase in grain protein concentration. The data for grain protein concentration during grain-filling (shown in figure 6.1) shows that under control temperatures protein concentration steadily decreases towards maturity. However, under elevated temperatures this initial decrease reverses at mid grain-filling, with the concentration of protein increasing again towards maturity. This could suggest a differential effect of elevated temperature on the later stages of grain-filling, where it may have a more detrimental effect on the accumulation of starch (or other cellular components). Whilst the post-anthesis temperature treatment used in this experiment resulted in changes to protein concentration, it did not significantly alter the protein *content*, *i.e.* the physical amount of protein within each grain, confirming previous reports (Koga *et al.* 2015). Therefore, it is clear that the observed changes in protein concentration with relation to temperature, are likely due to changes in starch accumulation. Nitrogen, however, did increase the amount (yield) of protein within each grain.

By comparing the nitrogen content data from the controlled-environment experiment to that collected from the WGIN diversity field experiment, it is possible to put into context the changes observed due to the nitrogen and temperature treatments used in this experiment. For the controlled-environment experiment, protein concentration at harvest was 8.5%/9.3% under the control temperature, and 12.7%/14.6% under the elevated temperature treatment for the low/high nitrogen treatments respectively. Cadenza grain collected from the field had a mean protein content of 8.0% under the low nitrogen treatment (100kg-N/ha), and 13.2% under the high nitrogen treatment (350kg-N/ha). Since the difference in grain nitrogen concentration between nitrogen treatments is much larger in the field experiment, is clear that the difference between the two nitrogen treatments used in the controlled-environment experiment was relatively small. Additionally, it illustrates how low the protein concentration of the grain grown under control temperatures is, with only the grain exposed to the elevated post-anthesis temperature treatment showing protein concentration comparable to that of grain grown in the field under sufficient nitrogen. This suggests that even plants grown under the high nitrogen treatment in the controlled-environment experiment were not supplied with sufficient nitrogen to achieve grain protein concentration at the level required for bread-making. The reason behind this observation may be that the amount of nitrogen provided to the plants by the high-nitrogen treatment was insufficient, or it may be a result of the fact that the controlled-environment experiment was a pot experiment, whereas the WGIN

diversity experiment was a field experiment.

SE-HPLC showed that both high nitrogen input, and elevated post-anthesis temperatures increased the gluten content of the mature grain, confirming previous reports (Malik *et al.* 2011; Moldestad *et al.* 2014; Tao *et al.* 2018). These results mirror the mature grain protein concentration results, and indicate that the observed increase in protein concentration also correlates with an increase in the proportion of gluten proteins. This is perhaps to be expected, since as storage proteins, the amount of gluten accumulated within the grain is inherently variable, whilst the albumin and globulin proteins are generally non-storage proteins associated with various cellular roles as enzymes or inhibitors etc. (Singh *et al.* 2001a), and may remain at a baseline level due to their importance in various cellular activities. Regardless, such increases in grain gluten content are associated with greater dough strength, extensibility, and elasticity, which results in improved bread-making performance (El Haddad *et al.* 1995; Sapirstein *et al.* 1998).

The canonical variate analysis of the SDS-PAGE data was performed as an initial investigatory step to predict and visualise the effect that both temperature and nitrogen treatments had on the protein composition throughout grain-filling. This analysis indicates that the protein composition of the mature grain through development was more greatly affected by elevated temperature than nitrogen input. This initial analysis was relatively accurate as a predictor of the final results, which generally showed a lack of significance with regards to the nitrogen treatment: of the 13 proteins measured, only four showed a significant response to nitrogen input alone. Analysis of the CV loadings identified proteins P3 and P4 as potentially more susceptible to the different temperature treatments. Whilst protein P3 did show a strong significant increase under the high nitrogen treatment, P4 showed no significant effect ($P=0.368$). This perhaps illustrates one of the limitations of CV analysis: its inability to account for data variability and limited replication in the same way as a technique such as ANOVA or REML.

Protein composition data was collected using both SDS-PAGE and SE-HPLC, which both measure different aspects of protein composition: whilst SDS-PAGE measures reduced protein subunits, SE-HPLC measures the amount of protein polymers present, grouped into one of five fractions enriched with a particular protein group. To make direct comparisons between the two datasets, the protein subunits identified by SDS-PAGE must be attributed to the same groups as those used in the SE-HPLC analysis. This allows us to directly compare the effect that each of the treatments used in the controlled-environment experiment had on each group of gluten proteins. When the two datasets are comparably grouped, it becomes apparent that there is little consensus between the two techniques. The only protein group that shows consistent responses to temperature is the HMW-glutenins. Elevated post-anthesis temperature increased the F1 fraction from the SE-HPLC analysis, and also increased the proportion of three of the four HMW-glutenin proteins identified by SDS-PAGE. However,

for the response to nitrogen, the two techniques differ: whilst SE-HPLC reported an increase in HMW-glutenin content under the high nitrogen treatment, SDS-PAGE generally reported a decrease. Elsewhere, the two techniques either contradict one another, or else show a lack of significant responses.

Whilst it is difficult to draw conclusions on the effect of elevated temperature and nitrogen supply on protein composition by analysing the relative levels of each protein fraction detected by SE-HPLC, the ratios of F1/F2 and (F3+F4)/F1 provide perhaps the most useful results from this analysis. Since these ratios are associated with key flour quality characteristics, the significant differences between these ratios are of particular interest. The plot of this data shown in figure 6.9 shows that grain grown under higher temperature during grain-filling produce flour better suited to bread-making (again confirming reports by Moldestad *et al.* (2014) and Tao *et al.* (2018)), and also that the effect of nitrogen input *prior* to anthesis is reduced when high temperatures are experienced during during grain-filling. This suggests that high nitrogen application rates may be less worthwhile in years where high temperatures are experienced during grain-filling in terms of achieving good bread-making quality.

6.4.2. WGIN diversity field trial experiment

Protein concentration data was collected over the three years of the WGIN diversity field experiment both during grain-filling, and at maturity. The data collected during grain-filling shows some interesting effects, and some comparisons can be made to the results of the controlled-environment experiment. In both of these experiments, increasing the nitrogen input increased grain protein concentration throughout grain-filling, an effect that persisted to maturity. Likewise in the controlled-environment experiment, the elevated temperature treatment also increase the relative grain protein concentration, an effect that is also present to a lesser degree between the years of the field experiment. Here, the difference between the coolest (2016) and hottest (2017) years are clear, particularly under the high nitrogen treatment, with the hottest year showing the highest protein concentration. As with the controlled-environment experiment, the effects observed during grain-filling continued to maturity, with the grain protein concentration of the mature grain being significantly increased by the high nitrogen treatment, and with the highest concentration of protein in the grain grown in the hottest year. A further consideration that can only be made from the field experiment results is that of genotype. The four different genotypes responded differently to the three years of the experiment. Whilst generally the lowest grain protein concentration measurements were recorded in 2016, for Cadenza they were recorded in 2015. Likewise, the year-to-year variation in grain protein concentration was lowest in Cadenza, and was highest in Hereward. This is an important consideration when growing bread-making wheat, since certain grain protein levels must be achieved consistently for the crop to be sold for bread-making purposes.

As in the controlled-environment experiment, increased nitrogen input lead to an increase in

the gluten content of the mature grain in the field experiment. Whilst a significant difference was not seen between years, different genotypes did have different gluten contents. Of the four genotypes, the group 1 bread-making wheat Hereward recorded the highest gluten content. Whilst this result may be predictable, the fact that Cadenza (a group two bread-making wheat) had a lower gluten content than feed wheat Istabraq is unexpected. However, it should be noted that the differences between genotypes were small, with all genotypes averaging 81–83% gluten content.

The exploratory CV analysis of the SDS-PAGE data suggested that the greatest effect on protein composition in the field experiment was from nitrogen treatment. Likewise, the analysis showed some differences between genotypes, showing that Hereward may be more susceptible to changes in protein composition due to nitrogen input that Soissons. The vector loadings from each protein group provided information on which protein groups may experience the greatest effects from changes in nitrogen input. Of the four protein groups, the LMW-glutenins and omega-gliadins showed the highest loading value for the CV that separated the two nitrogen treatment, suggesting that these protein groups were the most susceptible to nitrogen.

Unfortunately, for the SE-HPLC analysis of the WGIN diversity field experiment no significant effects were identified for the F1/F2 and (F3+F4)/F1 ratios. Furthermore, the analysis of the SE-HPLC protein composition data failed to show any significant effects on the proportion of either HMW- (F1) or LMW-glutenins (F2). This is likely due to inaccuracies in the quantification of the SE-HPLC plots, since the boundary between the peaks for F1 and F2 fractions is often impossible to accurately define, leading to incorrect measurements of F1 and F2 concentration. Whilst disappointing, this doesn't affect the quantification of the F3–F5 protein fractions, and so these results can still be interpreted.

The predictions made based on the results of the CV analysis were somewhat confirmed by the SDS-PAGE and SE-HPLC analysis of the grain protein composition. Nitrogen input and genotype had the greatest effect on protein composition, whilst an effect from year of experiment was only identified in the SE-HPLC measurement of omega-gliadins, whereby the hottest year correlated with the highest concentration of omega-gliadins. In the SDS-PAGE data, the high nitrogen treatment was found to decrease the concentration of LMW-glutenins, and increase the concentration of omega-gliadins, whilst in the SE-HPLC it was found to increase the proportion of alpha-, beta-, and gamma-gliadins. In each of these cases, the effect of nitrogen was greatest in Hereward, and smallest in Soissons, as predicted by the CV analysis. These results confirm previous findings: the concentration of gliadin proteins in the mature grain is increased, and the concentration of LMW-glutenins decreased by high nitrogen input (Chope *et al.* 2014).

Chapter 7: Gene expression analysis results

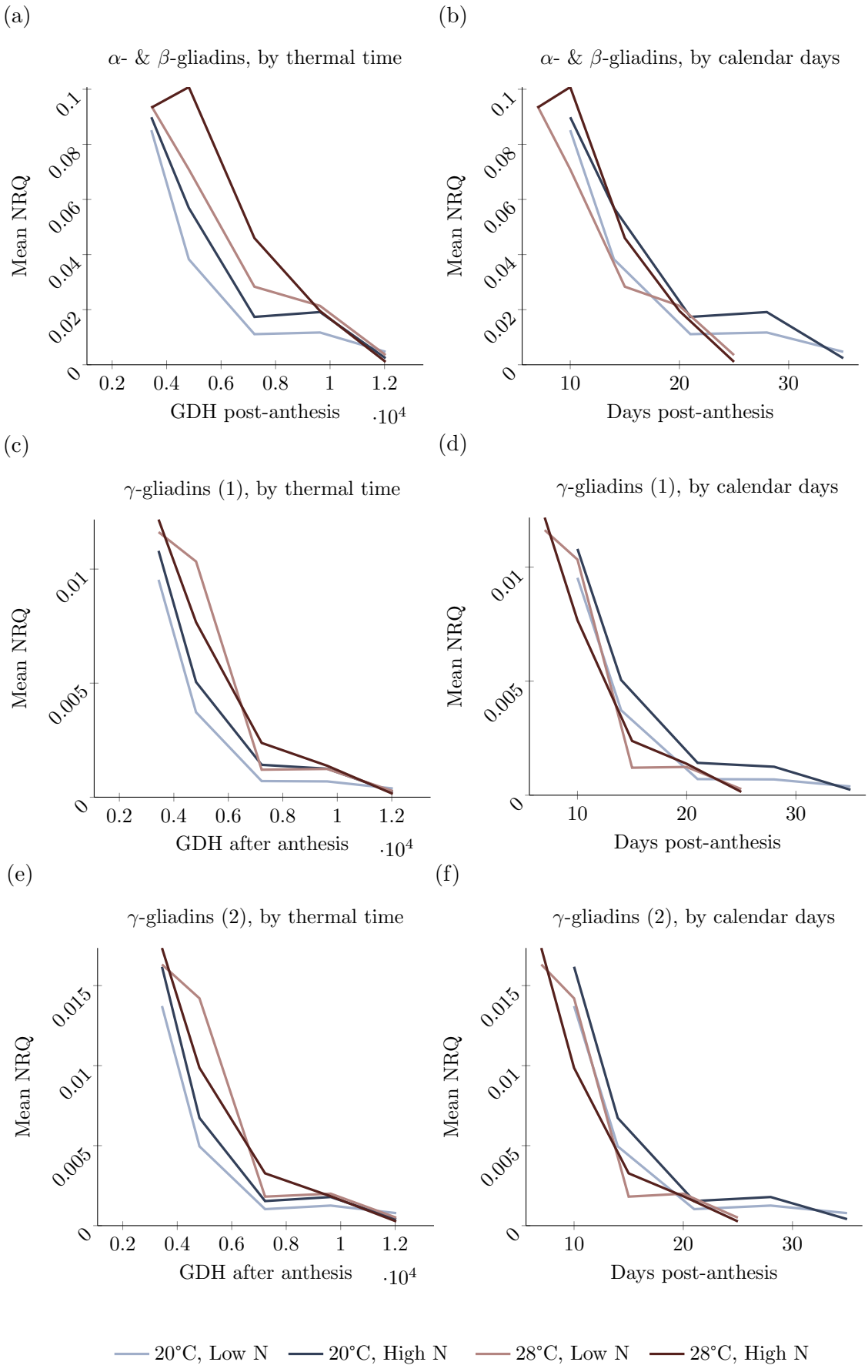
7.1. Introduction

Understanding the genetics responsible for crop quality is of great significance, both academically and commercially. By quantifying the effect that experimental factors have on gene expression we can identify the genetic mechanisms behind any physiological responses that are observed, information which can then be used to direct future research and breeding efforts. Whilst the use of genetic modification and gene-editing technologies are unlikely to be acceptable to consumers in Europe in the foreseeable future, knowledge of crop genetics can still be used in techniques such as marker-assisted breeding which combine an understanding of genetics with more traditional breeding techniques. In wheat, the primary determinant of quality is the protein composition and content within the grain (El Haddad *et al.* 1995; Sapirstein *et al.* 1998). Therefore understanding the factors that impact on the expression of wheat storage protein synthesis genes would be beneficial, both in terms of wider commercial and academic context. In therms of the present study, it is hoped that analysis of the expression of these protein synthesis genes will provide further understanding of the protein content, composition, and distribution results.

As part of both the controlled-environment and WGIN diversity field experiments, mRNA expression in spring-wheat variety Cadenza was analysed through qPCR of six storage protein synthesis gene transcripts: α- and β-gliadins, γ-gliadins (1), γ-gliadins (2), HMW-glutenins, LMW-glutenins, and ω-gliadins. These transcripts represent a complete coverage of the gluten storage proteins present in the wheat grain, with their expression recorded at five sampling timepoints during grain-filling. The aim of this analysis was to identify the effect that elevated temperature during grain-filling, nitrogen supply, and year-to-year variation in the field would have on the expression of storage protein synthesis genes, and also to identify any relationship between gene expression and protein composition (results presented in chapter 6) through linear regression analysis.

7.2. Controlled-environment experiment

In the controlled-environment experiment, British spring bread-making wheat Cadenza was supplied with either full or one-tenth strength nitrogen fertiliser, and was subjected to either a control (20°C) or elevated (28°C) daytime temperature treatment for the duration of grain-filling. Grain was sampled at five timepoints between anthesis and harvest for RNA expression analysis. The aim of this analysis was to identify the combined effect that nitrogen supply prior to anthesis and elevated temperature after anthesis have on the expression of storage protein synthesis genes over the course of grain-filling. The five sampling timepoints used in this experiment were adjusted for thermal time (see table 2.2) in an effort to collect samples at a comparable stage of development. Therefore each sampling timepoint (labelled T1–T5) represents a different number of calendar days after anthesis depending on the applied


(a)
α- & β-gliadins, by thermal time
Mean NRQ
0
0.02
0.04
0.06
0.08
0.1
0.2
0.4
0.6
0.8
1
1.2
GDH post-anthesis
·10⁴
(b)
α- & β-gliadins, by calendar days
Mean NRQ
10
20
30
Days post-anthesis
(c)
γ-gliadins (1), by thermal time
Mean NRQ
0
0.005
0.01
GDH after anthesis
·10⁴
(d)
γ-gliadins (1), by calendar days
Mean NRQ
Days post-anthesis
(e)
γ-gliadins (2), by thermal time
Mean NRQ
0.015
GDH after anthesis
·10⁴
(f)
γ-gliadins (2), by calendar days
Mean NRQ
Days post-anthesis
20°C, Low N
20°C, High N
28°C, Low N
28°C, High N

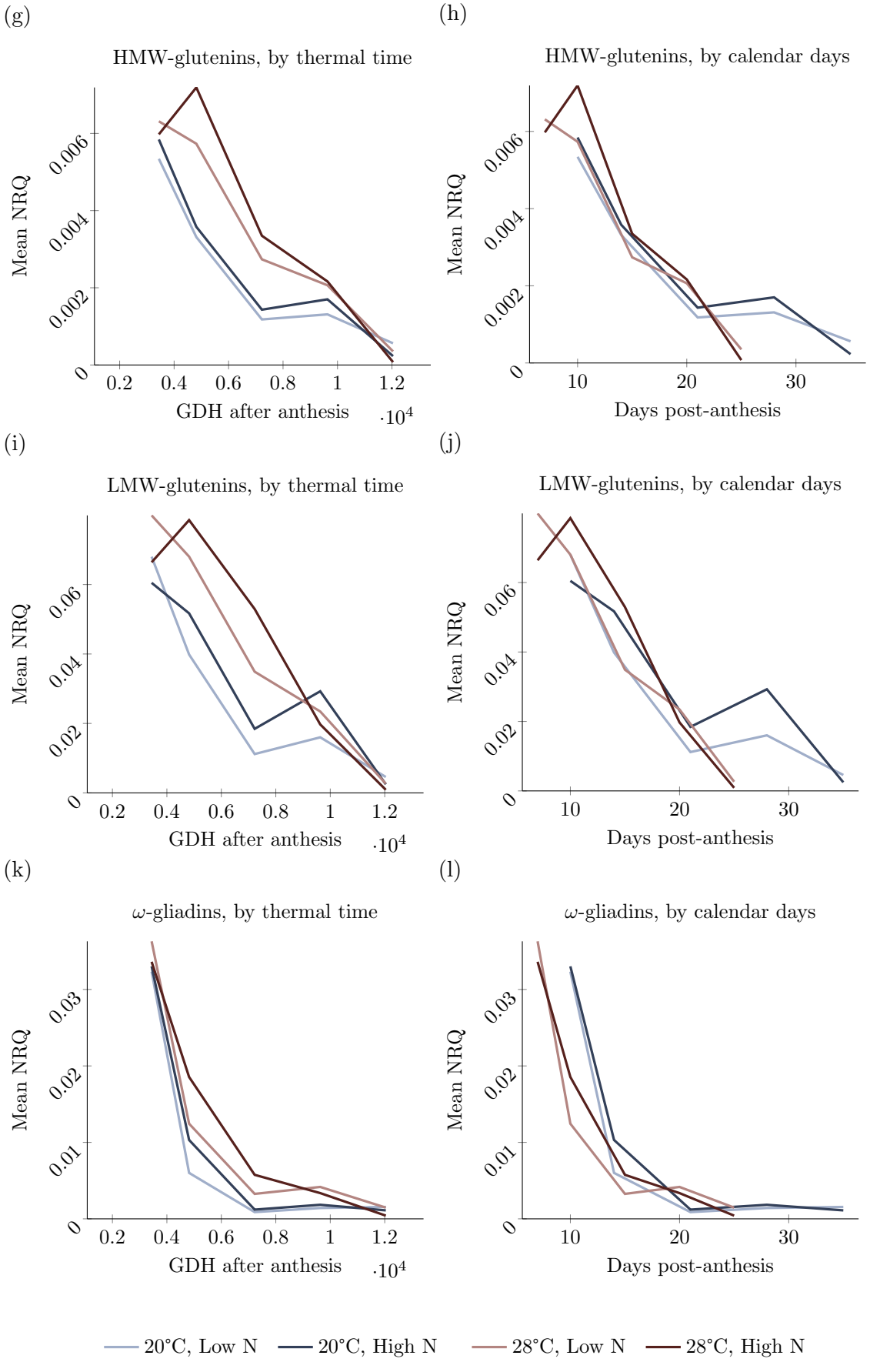

(g)
HMW-glutenins, by thermal time
Mean NRQ
0
0.002
0.004
0.006
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(h)
HMW-glutenins, by calendar days
Mean NRQ
0
0.002
0.004
0.006
10
20
30
Days post-anthesis
(i)
LMW-glutenins, by thermal time
Mean NRQ
0
0.02
0.04
0.06
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(j)
LMW-glutenins, by calendar days
Mean NRQ
0
0.02
0.04
0.06
10
20
30
Days post-anthesis
(k)
ω-gliadins, by thermal time
Mean NRQ
0
0.01
0.02
0.03
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(l)
ω-gliadins, by calendar days
Mean NRQ
0
0.01
0.02
0.03
10
20
30
Days post-anthesis
20°C, Low N
20°C, High N
28°C, Low N
28°C, High N

Figure 7.1: **Expression of wheat storage protein synthesis genes in Cadenza is differentially affected by elevated temperature over time**. Gene expression data showing the effect of nitrogen input and elevated post-anthesis temperature on the relative expression of six wheat grain storage protein synthesis genes over time in both calendar days after anthesis (DPA) and accumulated thermal time (GDH). Mean NRQ values for α- & β-gliadins by (a) thermal time and (b) calendar days, average standard error of the mean (SEM) 0.00903 (max 0.03417); γ-gliadins (1) by (c) thermal time and (d) calendar days, average SEM 0.00116 (max 0.00531); γ-gliadins (2) by (e) thermal time and (f) calendar days, average SEM 0.00169 (max 0.00805); HMW-glutenins by (g) thermal time and (h) calendar days, average SEM 0.00060 (max 0.00215); LMW-glutenins by (i) thermal time and (j) calendar days, average SEM 0.00776 (max 0.02459); and ω-gliadins by (k) thermal time and (l) calendar days, average SEM 0.00253 (max 0.01231). Analysis was completed on log-transformed data, but raw data is presented for clarity.

temperature treatment. For completeness, the RNA expression data was analysed against both accumulated thermal time (*i.e.* sampling timepoint) and days post anthesis (presented in figure 7.1). Analysing the data in this way reveals how gene expression is affected by temperature and nitrogen treatments both at a comparable level of maturity, and also at any point in time, regardless of developmental maturity. In addition to the raw data presented in figure 7.1, predictions from the appropriate REML models are also presented in tables 7.1–7.6, which allow for comparisons to be made between mean NRQ values on the log scale (as the data was analysed) with the relevant LSD at the 5% (p=0.05) level.

7.2.1. Alpha- and Beta-gliadins

The REML analysis of α- and β-gliadin gene expression data identified a significant interaction between both post-anthesis temperature treatment and sampling timepoint (as accumulated thermal time in GDH) squared ($F=12.94$, $P<0.001$) (figure 7.1a and table 7.1a); and between post-anthesis temperature treatment and calendar days post-anthesis squared ($F=19.14$, $P<0.001$) (figure 7.1b and table 7.1b). The effect of time (either in days or accumulated thermal time) was squared to investigate any non-linear effect over the course of grain-filling. The identification of these significant interactions indicates a non-linear trend between gene expression and time. No significant interaction was found with nitrogen treatment between either accumulated thermal time ($F=0.054$, $P=0.467$) or calendar days post-anthesis ($F=0.61$, $P=0.441$).

When analysed against sampling timepoints of equivalent accumulated thermal time, expression of the α- and β-gliadin synthesis gene transcript was increased by elevated temperature treatment during mid grain-filling (timepoint T3), but was decreased relative to the control temperature treatment towards the end of grain-filling (timepoint T5). In comparison, when analysed against calendar days after anthesis, elevated temperature did not result in an increase in α- and β-gliadin synthesis gene expression on any specific day during grain-

Table 7.1: Predicted mean NRQ values from the REML analysis of α- & β-gliadin synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-2.531	-3.104	-4.036	-4.879	-5.635
Elevated (28°C)	-2.693	-2.348	-2.856	-4.781	-8.123
Difference (* P≤0.05)	-0.162	0.756	1.180*	0.098	-2.488*
LSD	1.107	0.784	0.957	0.892	1.259
SED	0.548	0.388	0.474	0.442	0.623

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-2.083	-2.531	-3.104	-3.242	-3.908	-4.035	-4.528	-4.879	-5.635
Elevated (28°C)	-2.694	-2.337	-2.646	-2.863	-4.790	-5.371	-8.117	-10.786	-18.974
Difference (* P≤0.05)	-0.611	0.194	0.458	0.379	-0.882	-1.336*	-3.589*	-5.907*	-13.339*
LSD	1.434	0.960	0.857	0.874	0.914	0.915	1.151	1.704	4.214
SED	0.710	0.475	0.425	0.433	0.453	0.453	0.570	0.844	2.087

filling. However, as grain-filling progresses the elevated temperature treatment did result in a decrease in expression, which was significantly lower than that of the plants under the control-temperature treatment from 21DPA onwards.

7.2.2. Gamma-gliadins

Significant interactions were found between post-anthesis temperature treatment and sampling timepoint squared on the expression of both γ-gliadin transcript 1 (F=8.51, P=0.006) (figure 7.1c and table 7.2a) and 2 (F=9.30, P=0.004) (figure 7.1e and table 7.3a). Likewise, interactions were also identified between post-anthesis temperature and calendar days after anthesis for both transcript 1 (F=10.11, P=0.003) (figure 7.1d and table 7.2b) and 2 (F=9.95, P=0.003) (figure 7.1f and table 7.3b). No significant interaction was found between nitrogen treatment and accumulated thermal time for either transcript 1 (F=0.57, P=0.0456) or 2 (F=0.23, P=0.636), and no interaction was found between nitrogen treatment and calendar days after anthesis for either transcript 1 (F=0.91, P=0.346) or 2 (F=0.63, P=0.433).

The effect of elevated temperature on γ-gliadin protein synthesis gene expression was similar in each of the transcripts analysed against sampling timepoint: expression was decreased at the end of grain filling (T5). When analysed against calendar days after anthesis the results are exaggerated, with elevated temperature causing an increasing reduction in expression from 20DPA onwards.

Table 7.2: Predicted mean NRQ values from the REML analysis of γ-gliadin (1) synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-4.685	-5.498	-6.667	-7.514	-8.038
Elevated (28°C)	-4.682	-4.850	-5.790	-7.555	-10.144
Difference (* P≤0.05)	0.003	0.648	0.877	-0.041	-2.106*
LSD	1.036	0.728	0.893	0.831	1.180
SED	0.512	0.360	0.442	0.412	0.584

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-4.007	-4.685	-5.498	-5.685	-6.52	-6.664	-7.190	-7.514	-8.038
Elevated (28°C)	-4.679	-4.850	-5.538	-5.793	-7.557	-8.025	-10.143	-12.088	-17.778
Difference (* P≤0.05)	-0.672	-0.165	-0.040	-0.108	-1.037*	-1.361*	-2.953*	-4.574*	-9.740*
LSD	1.345	0.895	0.798	0.814	0.852	0.853	1.077	1.600	3.967
SED	0.666	0.443	0.395	0.403	0.422	0.422	0.533	0.792	1.964

Table 7.3: Predicted mean NRQ values from the REML analysis of γ-gliadin (2) synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-4.327	-5.202	-6.376	-7.097	-7.363
Elevated (28°C)	-4.374	-4.554	-5.465	-7.136	-9.568
Difference (* P≤0.05)	-0.047	0.648	0.911*	-0.039	-2.205*
LSD	1.054	0.744	0.910	0.848	1.200
SED	0.522	0.368	0.451	0.420	0.594

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-3.573	-4.327	-5.202	-5.397	-6.236	-6.372	-6.844	-7.097	-7.363
Elevated (28°C)	-4.370	-4.555	-5.225	-5.468	-7.138	-7.578	-9.566	-11.385	-16.691
Difference (* P≤0.05)	-0.797	-0.228	-0.023	-0.0710	-0.902*	-1.206*	-2.722*	-4.288*	-9.328*
LSD	1.367	0.912	0.814	0.830	0.869	0.869	1.096	1.625	4.024
SED	0.677	0.452	0.403	0.411	0.430	0.431	0.543	0.805	1.993

Table 7.4: Predicted mean NRQ values from the REML analysis of HMW-glutenin synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-5.307	-5.683	-6.368	-7.085	-7.836
Elevated (28°C)	-5.420	-4.925	-5.258	-7.119	-10.508
Difference (* P≤0.05)	-0.113	0.758*	1.110*	-0.034	-2.672*
LSD	1.046	0.704	0.888	0.820	1.202
SED	0.518	0.348	0.440	0.406	0.595

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-5.032	-5.307	-5.683	-5.779	-6.268	-6.368	-6.774	-7.085	-7.836
Elevated (28°C)	–5.423	-4.911	-5.074	-5.266	-7.129	-7.713	-10.501	-13.248	-21.770
Difference (* P≤0.05)	-0.391	0.396	0.609	0.513	-0.861*	-1.345*	-3.727*	-6.163*	-13.934*
LSD	1.382	0.892	0.783	0.801	0.844	0.845	1.090	1.653	4.154
SED	0.684	0.442	0.388	0.397	0.418	0.418	0.540	0.818	2.057

7.2.3. HMW-glutenins

The expression of HMW-glutenin synthesis gene transcripts was significantly affected by an interaction between both elevated temperature treatment and sampling timepoint squared (F=13.05, P<0.001) (figure 7.1g and table 7.4a), and between elevated temperature and calendar days after anthesis squared (F=20.82, P<0.001) (figure 7.1h and table 7.4b). No significant interaction was found between either nitrogen treatment and accumulated thermal time (F=0.31, P=0.583) or between nitrogen and calendar days after anthesis (F=0.38, P=0.539).

In the analysis against sampling timepoint, elevated temperature resulted in an increase in the expression of HMW-glutenin synthesis gene transcripts, at T2 and T3, but a reduction in expression at the final sampling timepoint (T5). When analysed against calendar days after anthesis, elevated temperature did not result in a significant increase in expression on any specific day, but again resulted in a decrease in expression that increased in effect from 20DPA onwards.

7.2.4. LMW-glutenins

A significant interaction between temperature treatment and sampling timepoint squared (F=13.38, F<0.001) (figure 7.1i and table 7.5a) and between temperature and calendar days after anthesis squared (F=22.30, P<0.001) (figure 7.1j and table 7.5b) was found on the

Table 7.5: Predicted mean NRQ values from the REML analysis of LMW-glutenin synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-2.827	-3.173	-3.864	-4.667	-5.580
Elevated (28°C)	-2.979	-2.401	-2.724	-4.744	-8.462
Difference (* P≤0.05)	-0.152	0.772	1.140*	-0.077	-2.882*
LSD	1.132	0.789	0.973	0.905	1.292
SED	0.561	0.391	0.482	0.448	0.640

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-2.592	-2.827	-3.173	-3.265	-3.759	-3.866	-4.309	-4.667	-5.580
Elevated (28°C)	-2.982	-2.386	-2.530	-2.733	-4.756	-5.395	-8.455	-11.478	-20.878
Difference (* P≤0.05)	-0.390	0.441	0.643	0.532	-0.997*	-1.529*	-4.146*	-6.811*	-15.298*
LSD	1.477	0.977	0.868	0.886	0.929	0.929	1.179	1.758	4.372
SED	0.731	0.484	0.430	0.439	0.460	0.460	0.584	0.871	2.165

expression of LMW-glutenin transcripts. Again, there was no significant effect from the combination of nitrogen treatment and sampling timepoint (F=0.57, P=0.455) or between nitrogen and days after anthesis (F=0.68, P=0.414).

The elevated temperature treatment resulted in an increase in expression of the LMW-glutenin synthesis transcript at sampling timepoint T3, and a decrease in expression at timepoint T5. As with previous transcripts, when analysed against calendar days after anthesis there was no significant increase in expression on any specific day attributed to the elevated temperature treatment, only an increasing reduction in expression from 20DPA onwards.

7.2.5. Omega-gliadins

The expression of ω-gliadin synthesis gene transcripts was significantly affected by an interaction between post-anthesis temperature treatment and sampling timepoint squared (F=19.11, F<0.001) (figure 7.1k and table 7.6a), and by and interaction between temperature treatment and calendar days after anthesis squared (F=12.40, F<0.001) (figure 7.1l and table 7.6b). There was no significant interaction between nitrogen treatment and timepoint (F=0.38, P<0.001) or nitrogen treatment and calendar days after anthesis (F=0.15, P=0.705).

In the REML analysis of ω-gliadin synthesis gene expression with sampling timepoint, elevated temperature resulted in an increase in expression at timepoints T2 and T3, and a decrease in expression at timepoint T5. When analysed against calendar days after anthesis, expression

Table 7.6: Predicted mean NRQ values from the REML analysis of ω-gliadin synthesis gene expression data from the controlled-environment experiment, presented on the log scale. (a) the interaction between post-anthesis temperature treatment and sampling timepoint (or accumulated GDH after anthesis) squared; and (b) the interaction between post-anthesis temperature treatment and calendar days after anthesis squared. LSD (at the 5% level) and SED values are presented.

(a)

	Sampling timepoint (accumulated GDH after anthesis)				
Temperature treatment	T1 (3440)	T2 (4816)	T3 (7224)	T4 (9632)	T5 (12040)
Control (20°C)	-3.507	-4.959	-6.608	-7.120	-6.497
Elevated (28°C)	-3.695	-4.068	-5.036	-6.405	-8.176
Difference (* P≤0.05)	-0.188	0.891*	1.572*	0.715	-1.679*
LSD	0.940	0.675	0.817	0.764	1.066
SED	0.466	0.334	0.404	0.378	0.528

(b)

	Calendar days after anthesis								
Temperature treatment	7	10	14	15	20	21	25	28	35
Control (20°C)	-2.175	-3.507	-4.959	-5.264	-6.442	-6.596	-7.040	-7.120	-6.497
Elevated (28°C)	-3.687	-4.073	-4.814	-5.039	-6.406	-6.736	-8.174	-9.428	-12.916
Difference (* P≤0.05)	-1.512*	-0.566	0.145	0.225	0.036	-0.140	-1.134*	-2.308*	-6.419*
LSD	1.209	0.818	0.735	0.748	0.781	0.782	0.976	1.434	3.526
SED	0.599	0.405	0.364	0.371	0.387	0.387	0.483	0.710	1.746

was not increased on any specific day by the elevated temperature treatment, but was decreased at 7DPA, and also from 25DPA onwards.

7.2.6. Comparison with SE-HPLC data

To relate the observations in gene expression data to the amount of protein present in the mature grain, a linear regression analysis was completed to correlate accumulated NRQ during grain-filling with the absolute protein content measurements from the relevant SE-

Table 7.7: Linear regression analysis comparing protein synthesis gene expression to SE-HPLC protein fractions in the controlled-environment experiment. The coefficient of determination (R^2) and P-value from the regression analysis of accumulated NRQ value over all five sampling timepoints and the absolute protein fraction measurements from the SE-HPLC analysis of mature grain. SE-HPLC fraction F1 is enriched with HMW-glutenins, F2 with LMW-glutenins, F3 with ω-gliadins, and F4 with α-, β-, and γ-gliadins. For the two γ-gliadin transcripts analysed, the mean value of both transcripts was used.

Gene transcript(s)	SE-HPLC fraction	R^2	P value
HMW-glutenins	F1	0.546	0.006
LMW-glutenins	F2	0.523	0.008
ω-gliadins	F3	0.332	0.050
α-, β-gliadins and γ-gliadins(1&2)	F4	0.549	0.006

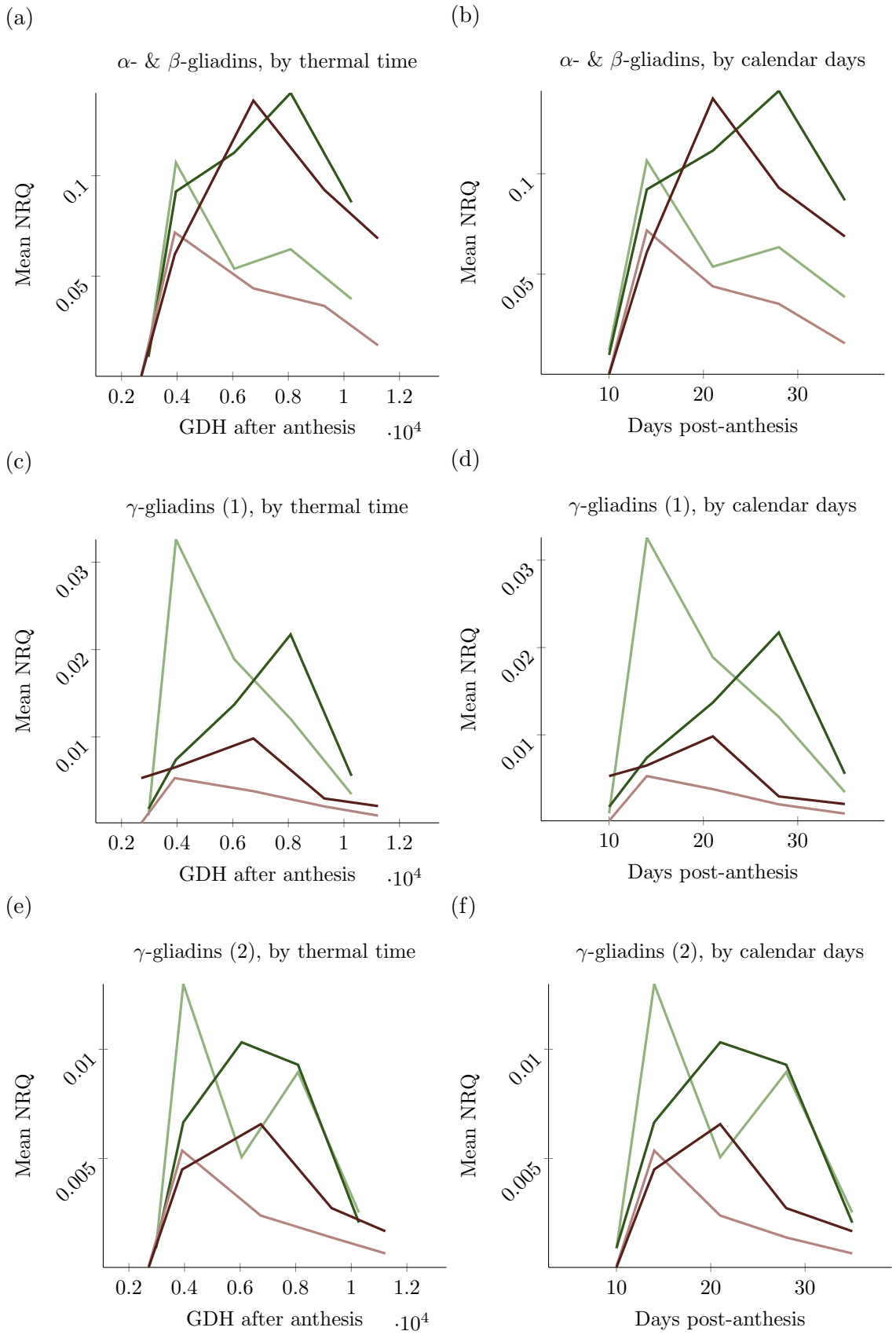
HPLC fraction. This analysis made no consideration of the different combinations of nitrogen and post-anthesis temperature, and was run solely to verify a link between protein synthesis gene expression, and the amount of the relevant protein in the mature grain. The coefficient of determination (R^2) and P-values from this analysis are presented in table 7.7. This analysis found that the expression of HMW-glutenin, LMW-glutenin, and α-, β- and γ-gliadin synthesis genes is tightly associated with the quantity of the relevant SE-HPLC protein fraction detected in the mature grain. The expression of the ω-gliadins synthesis gene transcript was more weakly associated with the SE-HPLC data, and was borderline significant at the 5% level.

7.3. **WGIN diversity field trial experiment**

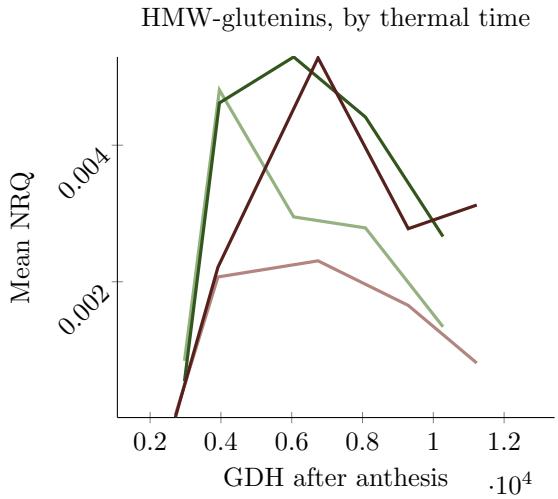
As part of the WGIN diversity field trial experiment, developing grain was sampled from Cadenza wheat plants grown under either 100kg-N/ha or 350kg-N/ha applied nitrogen in both 2016 and 2017. These two years represent years with considerable differences in temperature and sunlight (as discussed in chapter 3) in an effort to identify any year-to-year variation in storage protein synthesis gene expression, and how this is affected by nitrogen supply. Analysis was completed on Cadenza to facilitate comparisons with the controlled-environment experiment. Grain was sampled at five timepoints through grain-filling (see section 2.3.4), with no adjustment made for the differences in accumulated thermal time between years. Whilst not as significant as the differences between temperature treatments in the controlled-environment experiment, there was still a considerable difference in the thermal time accumulated by 35DPA in 2016 and 2017 (10133GDH and 11608GDH respectively). Therefore, as in the data presented from the controlled-environment experiment in section 7.2, the gene expression data from the WGIN field experiment will be presented in terms of both sampling timepoint (in this instance calendar days after anthesis) and accumulated thermal time (in GDH). The raw gene expression data from the WGIN field experiment is presented in figure 7.2, and predictions from the appropriate REML models are presented with the relevant LSD at the 5% (p=0.05) level in tables 7.8–7.12.

7.3.1. Alpha- and Beta-gliadins

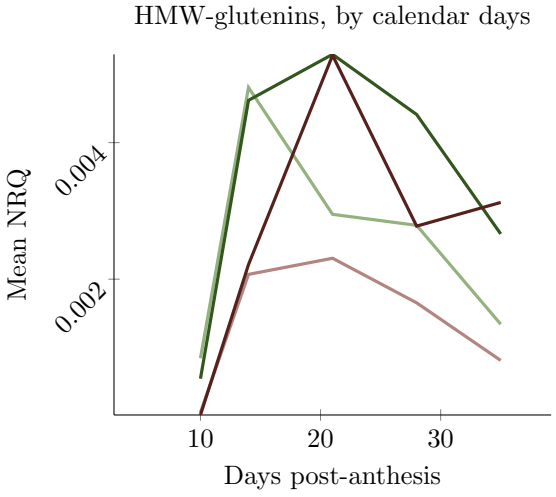
Analysis of the α- and β-gliadin gene expression data from the WGIN diversity field trial experiment with REML identified a significant interaction between accumulated thermal time (in GDH) and nitrogen treatment (F=5.07, P=0.029) (table 7.8a), and between accumulated thermal time and year of experiment (F=11.02, P=0.002) (table 7.8a), with the raw data behind each interaction presented in figure 7.2a. No interaction was found between nitrogen input, year of experiment, and accumulated thermal time (F=0.81, P=0.371). When analysed against sampling timepoint (or calendar days after anthesis), an interaction was identified between sampling timepoint and nitrogen treatment (F=5.20, P=0.027) (table 7.8c), and between sampling timepoint squared (indication a non-linear trend) and year of experiment (F=9.37, P=0.004) (table 7.8d). The raw data for both of these interactions are presented in figure 7.2b. Again, no interaction was found between nitrogen input, year of experiment, and sampling timepoint (F=0.57, P=0.455).


(a)
α- & β-gliadins, by thermal time
Mean NRQ
0.1
0.05
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(b)
α- & β-gliadins, by calendar days
Mean NRQ
0.1
0.05
10
20
30
Days post-anthesis
(c)
γ-gliadins (1), by thermal time
Mean NRQ
0.03
0.02
0.01
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(d)
γ-gliadins (1), by calendar days
Mean NRQ
0.03
0.02
0.01
10
20
30
Days post-anthesis
(e)
γ-gliadins (2), by thermal time
Mean NRQ
0.01
0.005
0.2
0.4
0.6
0.8
1
1.2
GDH after anthesis
·10⁴
(f)
γ-gliadins (2), by calendar days
Mean NRQ
0.01
0.005
10
20
30
Days post-anthesis
2017, 100kg-N/ha
2017, 350kg-N/ha
2016, 100kg-N/ha
2016, 350kg-N/ha

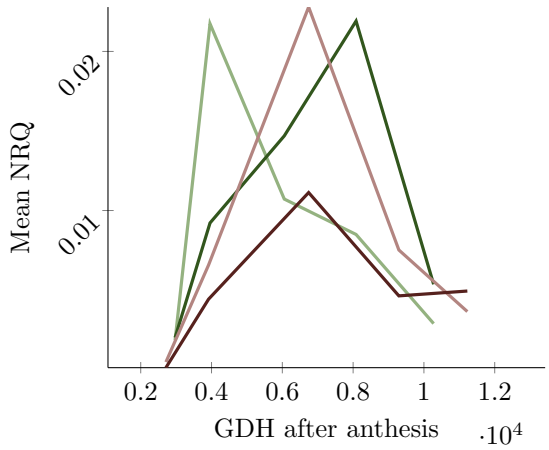
(g)



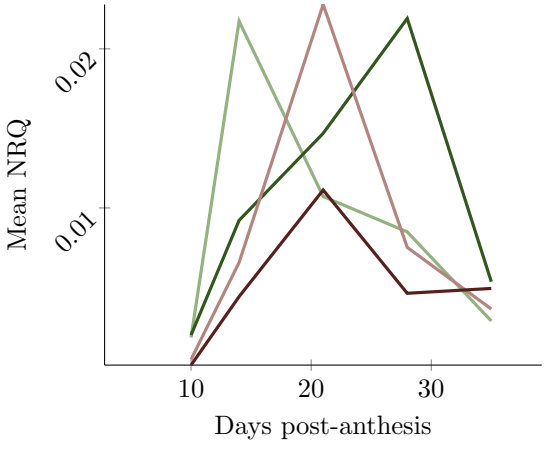
(h)



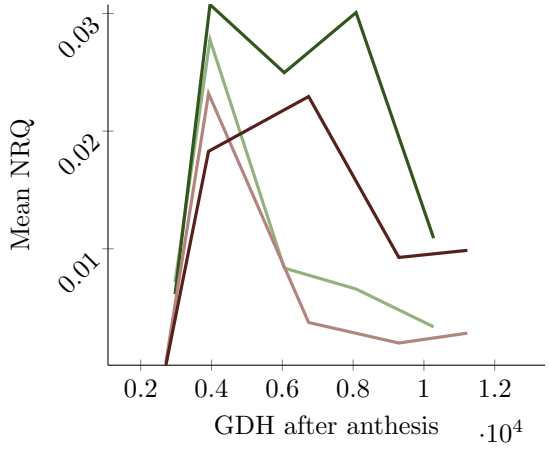
(i)



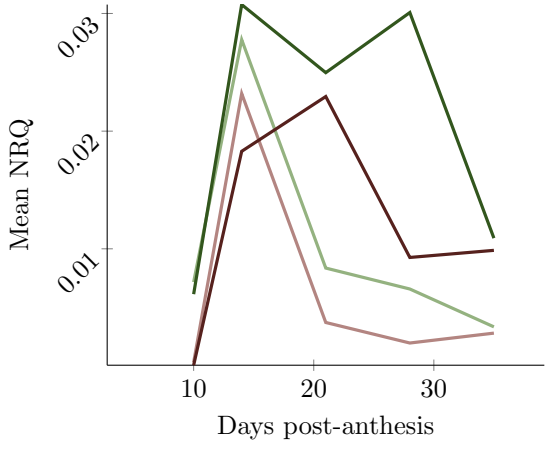
(j)



(k)



(l)



2017, 100kg-N/ha — 2017, 350kg-N/ha — 2016, 100kg-N/ha — 2016, 350kg-N/ha

Figure 7.2: **Expression of wheat storage protein synthesis genes in Cadenza is affected by nitrogen input and varies between years in the field**. Gene expression data showing the effect of nitrogen input on spring-wheat variety Cadenza over two years of the WGIN diversity field trial experiment on the relative expression of six wheat grain storage protein synthesis genes over time in both calendar days (DPA) and accumulated thermal time (GDH). Mean NRQ values for α- & β-gliadins by (a) thermal time and (b) calendar days, average SEM 0.01458 (max 0.05159); γ-gliadins (1) by (c) thermal time and (d) calendar days, average SEM 0.03638 (max 0.02669); γ-gliadins (2) by (e) thermal time and (f) calendar days, average SEM 0.00164 (max 0.00870); HMW-glutenins by (g) thermal time and (h) calendar days, average SEM 0.00060 (max 0.00174); LMW-glutenins by (i) thermal time and (j) calendar days, average SEM 0.00370 (max 0.02034); and ω-gliadins by (k) thermal time and (l) calendar days, average SEM 0.00279 (max 0.01180). Analysis was completed on log-transformed data, with raw data presented for clarity.

Regardless of the analysis approach taken (regressing against accumulated time versus against calendar days after anthesis), the results are comparable. Comparison of the predicted means representing the weakly significant effect from nitrogen treatment show that when compared on a like-for-like basis (either the same sampling timepoint or the same amount of accumulated thermal time), the difference in gene expression between the two nitrogen treatments isn't significant at the 5% level. However, the effect from year of experiment is stronger, with a linear trend against accumulated thermal time which is characterised by 2017 showing reduced expression during early grain-filling (T1–T2), but increased expression at the end of grain-filling (T5), when compared to 2016. Whilst the results from the non-linear trend against sampling timepoint also detected a significant decrease in expression during early grain-filling in 2017, there was no significant increase towards the end of grain-filling.

7.3.2. Gamma-gliadins

The REML analysis of γ-gliadin (2) synthesis gene expression identified a significant interaction between accumulated thermal time and year of experiment (F=7.36, P=0.009) (table 7.9a and figure 7.2e) and also between sampling timepoint squared and year of experiment (F=4.43, P=0.041) (table 7.9b and figure 7.2f). No significant interaction was found between either nitrogen treatment and accumulated thermal time (F=2.33, P=0.133) or between nitrogen and sampling timepoint (F=2.15, P=0.149). Conversely, the analysis of expression of the γ-gliadin (1) transcript failed to identify a significant effect between either nitrogen (F=2.49, P=0.121) or year (F=3.24, P=0.078) and accumulated thermal time (figure 7.2c); nor between nitrogen (F=2.34, P=0.133) or year (F=1.32, P=0.256) and sampling timepoint (figure 7.2d).

The comparison of the predicted means from the non-linear relationship between sampling and γ-gliadin (2) synthesis gene expression shows that expression was significantly lower in 2017 during early grain-filling (T1–T2), but wasn't significantly different at any other timepoint.

Table 7.8: Predicted mean NRQ values from the REML analysis of α- & β-gliadin synthesis gene expression from the WGIN diversity field trial experiment, presented on the log scale. The interaction between accumulated thermal time (GDH) after anthesis and (a) nitrogen treatment, (b) year of experiment; (c the interaction between sampling timepoint (or calendar days after anthesis) and nitrogen treatment; and (d) the interaction between sampling timepoint squared and year of experiment. LSD (at the 5% level) and SED values are presented.

(a)

	Accumulated GDH after anthesis				
Nitrogen treatment	2712	3916	6744	9292	11227
100kg-N/ha	-10.276	-8.202	-3.337	1.049	4.378
350kg-N/ha	-13.930	-10.582	-2.725	4.358	9.734
Difference (* P≤0.05)	-3.654	-2.380	0.612	3.309	5.356
LSD	5.669	3.922	0.734	4.072	6.883
SED	2.807	1.942	0.363	2.016	3.408

(b)

	Accumulated GDH after anthesis				
Year of experiment	2712	3916	6744	9292	11227
2016	-8.225	-6.597	-2.775	0.670	3.284
2017	-15.980	-12.187	-3.286	4.738	10.828
Difference (* P≤0.05)	-7.755*	-5.590*	-0.511	4.068	7.544*
LSD	5.669	3.922	0.734	4.072	6.883
SED	2.807	1.942	0.363	2.016	3.408

(c)

	Sampling timepoint (calendar days after anthesis)				
Nitrogen treatment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
100kg-N/ha	10.940	-8.456	-4.110	0.236	4.583
350kg-N/ha	-14.753	-10.828	-3.960	2.908	9.777
Difference (* P≤0.05)	-3.813	-2.372	0.150	2.672	5.194
LSD	5.257	3.477	0.685	2.948	6.062
SED	2.603	1.722	0.339	1.460	3.002

(d)

	Sampling timepoint (calendar days after anthesis)				
Year of experiment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
2016	-4.034	-3.177	-2.325	-2.273	-3.043
2017	-7.433	-4.922	-2.350	-2.018	-3.999
Difference (* P≤0.05)	-3.399*	-1.745*	-0.025	0.255	-0.956
LSD	1.238	0.830	1.016	0.927	1.347
SED	0.613	0.411	0.503	0.459	0.667

Table 7.9: Predicted mean NRQ values from the REML analysis of γ-gliadin (2) synthesis gene expression from the WGIN diversity field trial experiment, presented on the log scale. (a) the interaction between year of experiment and accumulated thermal time (GDH); and (b) the interaction between year of experiment and sampling timepoint (or calendar days after anthesis). LSD (at the 5% level) and SED values are presented.

(a)

	Accumulated GDH after anthesis				
Year of experiment	2712	3916	6744	9292	11227
2016	-12.206	-10.190	-5.459	-1.194	2.043
2017	-17.493	-14.122	-6.212	0.920	6.332
Difference (* P≤0.05)	-5.287	-3.932*	-0.753*	2.114	4.289
LSD	5.590	3.867	0.724	4.015	6.787
SED	2.768	1.915	0.358	1.988	3.361

(b)

	Sampling timepoint (calendar days after anthesis)				
Year of experiment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
2016	-6.647	-5.663	-4.793	-4.970	-6.227
2017	-9.540	-7.414	-5.353	-5.334	-7.421
Difference (* P≤0.05)	-2.893*	-1.751*	-0.560	-0.364	-1.194
LSD	1.241	0.832	1.018	0.929	1.350
SED	0.614	0.412	0.504	0.460	0.668

However when this data is analysed against accumulated thermal time, the level of expression in 2017 was significantly lower during early to mid grain-filling (approximately T2–T3), but not at the start of grain-filling (approximately T1).

7.3.3. HMW-glutenins

The expression of HMW-glutenin synthesis genes was significantly affected by interactions between accumulated thermal time and both nitrogen (F=4.35, P=0.042) (table 7.10a), and year of experiment (F=15.28, P<0.001) (table 7.10b, data for both interactions presented in figure 7.2g); and also by interactions between sampling timepoint and nitrogen treatment (F=4.42, P=0.041) (table (7.10c) and sampling timepoint *squared* and year of experiment (F=6.12, P=0.017) (table (7.10d, data for both interactions presented in figure 7.2h). There were no significant interactions found between either accumulated thermal time, nitrogen treatment, and year of experiment (F=1.47, P=0.231), or between sampling timepoint, nitrogen, and year (F=1.77, 0.189).

Comparison of the predicted mean NRQ values for the interactions between both nitrogen treatment and accumulated thermal time, and nitrogen and sampling timepoint found that although there was a significant overall effect of nitrogen, when there were no significant differences between the two treatments when comparing at the same point in time (either in terms of accumulated thermal time of calendar days after anthesis). However, the effect from year of experiment was greater, with HMW-glutenin synthesis gene expression higher

Table 7.10: Predicted mean NRQ values from the REML analysis of HMW-glutenin synthesis gene expression from the WGIN diversity field trial experiment, presented on the log scale. The interaction between accumulated thermal time (GDH) after anthesis and (a) nitrogen treatment, (b) year of experiment; (c the interaction between sampling timepoint (or calendar days after anthesis) and nitrogen treatment; and (d) the interaction between sampling timepoint squared and year of experiment.. LSD (at the 5% level) and SED values are presented.

(a)

	Accumulated GDH after anthesis				
Nitrogen treatment	2712	3916	6744	9292	11227
100kg-N/ha	-13.486	-11.361	-6.373	-1.877	1.536
350kg-N/ha	-15.626	-12.736	-5.955	0.159	4.799
Difference (* P≤0.05)	-2.140	-1.375	0.418	2.036	3.263
LSD	4.629	3.203	0.599	3.325	5.621
SED	2.292	1.586	0.297	1.646	2.783

(b)

	Accumulated GDH after anthesis				
Year of experiment	2712	3916	6744	9292	11227
2016	-11.986	-10.181	-5.947	-2.131	0.766
2017	-17.127	-13.916	-6.381	0.413	5.569
Difference (* P≤0.05)	-5.141*	-3.735*	-0.434	2.544	4.803
LSD	4.629	3.203	0.599	3.325	5.621
SED	2.292	1.586	0.297	1.646	2.783

(c)

	Sampling timepoint (calendar days after anthesis)				
Nitrogen treatment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
100kg-N/ha	-13.904	-11.440	-7.126	-2.813	1.500
350kg-N/ha	-16.227	-12.867	-6.986	-1.105	4.776
Difference (* P≤0.05)	-2.323	-1.427	0.140	1.708	3.276
LSD	4.293	2.839	0.559	2.407	4.950
SED	2.126	1.406	0.277	1.192	2.451

(d)

	Sampling timepoint (calendar days after anthesis)				
Year of experiment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
2016	-7.135	-6.226	-5.384	-5.464	-6.495
2017	-9.909	-7.790	-5.605	-5.295	-6.919
Difference (* P≤0.05)	-2.774*	-1.564*	-0.221	0.169	-0.424
LSD	1.011	0.678	0.830	0.757	1.100
SED	0.501	0.336	0.411	0.375	0.545

in 2016 during early grain-filling (T1–T2) in both the thermal time and sampling timepoint analyses.

7.3.4. LMW-glutenins

The expression of LMW-glutenin synthesis was significantly affected by an interaction between year and accumulated thermal time (F=11.63, P=0.001) (table 7.11a and figure 7.2i), and also by an interaction between year of experiment and sampling timepoint (F=7.80, P=0.007) (table 7.11b and figure 7.2j). There was no significant interaction between either nitrogen treatment and accumulated thermal time (F=2.19, P=0.146) or nitrogen and sampling timepoint (F=2.47, P=0.123).

The effect that year of experiment had on the expression of LMW-glutenin synthesis genes was a reduction in expression in early to mid grain-filling (T1–T3) in 2017 compared to 2016. However, this effect was only present in the analysis against sampling timepoint, with any comparisons made between the nitrogen treatments at a specific accumulated thermal time not showing significance at the 5% level.

7.3.5. Omega-gliadins

The expression of ω-gliadin synthesis genes was found to be significantly effected by both accumulated thermal time and nitrogen treatment (F=6.23, P=0.016) (table 7.12a), and accumulated time and year of experiment (F=11.26, P=0.002) (table 7.12b, and figure 7.2k

Table 7.11: Predicted mean NRQ values from the REML analysis of LMW-glutenin synthesis gene expression from the WGIN diversity field trial experiment, presented on the log scale. (a) the interaction between year of experiment and accumulated thermal time (GDH); and (b) the interaction between year of experiment and sampling timepoint (or calendar days after anthesis). LSD (at the 5% level) and SED values are presented.

(a)

	Accumulated GDH after anthesis				
Year of experiment	2712	3916	6744	9292	11227
2016	-11.073	-9.240	-4.938	-1.059	1.885
2017	-14.496	-11.767	-5.362	0.411	4.793
Difference (* P≤0.05)	-3.423	-2.527	-0.424	1.470	2.908
LSD	4.760	3.310	0.769	3.434	5.771
SED	2.357	1.639	0.381	1.701	2.857

(b)

	Sampling timepoint (calendar days after anthesis)				
Year of experiment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
2016	-10.441	-8.604	-5.389	-2.174	1.041
2017	-15.757	-12.387	-6.489	-0.591	5.307
Difference (* P≤0.05)	-5.316*	-3.783*	-1.100*	1.583	4.266
LSD	4.508	3.003	0.756	2.559	5.191
SED	2.232	1.487	0.374	1.267	2.570

Table 7.12: Predicted mean NRQ values from the REML analysis of ω-gliadin synthesis gene expression from the WGIN diversity field trial experiment, presented on the log scale. The interaction between accumulated thermal time (GDH) after anthesis and (a) nitrogen treatment, (b) year of experiment; and the interaction between sampling timepoint (or calendar days after anthesis) and (c) nitrogen treatment, (d) year of experiment. LSD (at the 5% level) and SED values are presented.

(a)

	Accumulated GDH after anthesis				
Nitrogen treatmentr	2712	3916	6744	9292	11227
100kg-N/ha	-8.413	-7.483	-5.302	-3.336	-1.844
350kg-N/ha	-13.008	-10.406	-4.299	1.206	5.385
Difference (* P≤0.05)	-4.595	-2.923	1.003*	4.542*	7.229*
LSD	5.533	3.828	0.716	3.974	6.717
SED	2.740	1.895	0.355	1.968	3.326

(b)

	Accumulated GDH after anthesis				
Year of experiment	2712	3916	6744	9292	11227
2016	-7.901	-6.876	-4.470	-2.301	-0.655
2017	-13.520	-11.013	-5.131	0.172	4.196
Difference (* P≤0.05)	-5.619*	-4.137*	-0.661	2.473	4.851
LSD	5.533	3.828	0.716	3.974	6.717
SED	2.740	1.895	0.355	1.968	3.326

(c)

	Sampling timepoint (calendar days after anthesis)				
Nitrogen treatment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
100kg-N/ha	-8.837	-7.672	-5.634	-3.595	-1.557
350kg-N/ha	-13.595	-10.557	-5.240	0.077	5.393
Difference (* P≤0.05)	-4.758	-2.885	0.394	3.672*	6.950*
LSD	5.263	3.481	0.686	2.951	6.069
SED	2.606	1.724	0.340	1.461	3.005

(d)

	Sampling timepoint (calendar days after anthesis)				
Year of experiment	T1 (10)	T2 (14)	T3 (21)	T4 (28)	T5 (35)
2016	-7.676	-6.591	-4.692	-2.793	-0.893
2017	-14.756	-11.638	-6.182	-0.726	4.730
Difference (* P≤0.05)	-7.080*	-5.047*	-1.490*	2.067	5.623
LSD	5.263	3.481	0.686	2.951	6.069
SED	2.606	1.724	0.340	1.461	3.005

Table 7.13: Linear regression analysis comparing protein synthesis gene expression in Cadenza to SE-HPLC protein fractions in the WGIN diversity field trial experiment experiment. The coefficient of determination (R^2) and P-value from the regression analysis of accumulated NRQ value over all five sampling timepoints and the absolute protein fraction measurements from the SE-HPLC analysis of mature grain. SE-HPLC fraction F1 is enriched with HMW-glutenins, F2 with LMW-glutenins, F3 with ω-gliadins, and F4 with α-, β-, and γ-gliadins. For the two γ-gliadin transcripts analysed, the mean value of both transcripts was used.

Gene transcript(s)	SE-HPLC fraction	R^2	P value
HMW-glutenins	F1	0.1573	0.202
LMW-glutenins	F2	0.0619	0.435
ω-gliadins	F3	0.0004	0.949
α-, β-gliadins and γ-gliadins(1&2)	F4	0.1047	0.304

for both interactions); as well as by sampling timepoint and nitrogen (F=6.15, P=0.017) (table 7.12c), and sampling timepoint and year (F=8.53, P=0.005) (table 7.12d, and figure 7.2l for both interactions). There were no significant interactions between nitrogen treatment, year of experiment, and accumulated thermal time (F=0.69, P=0.411) or sampling timepoint (F=0.56, P=0.458).

The effect of high (350kg-N/ha) nitrogen input was an increase in the relative expression of the ω-gliadin synthesis gene transcript towards mid to late grain-filling (T4–T5), with comparable results when expressed against either accumulated thermal time or sampling timepoint. Likewise, the effect of year of experiment was comparable against both the thermal time and the sampling timepoint analysis, with a significant reduction in expression in early to mid grain-filling (T1–T3) in 2017 compared to 2016.

7.3.6. Comparison with SE-HPLC data

As in the controlled-environment experiment, a linear regression analysis was completed on the gene expression and SE-HPLC data from the WGIN diversity field trial experiment, to identify any correlation between protein synthesis gene expression and the amount of the relevant protein present in the grain at maturity. This analysis was completed on the dataset at a whole, with no factoring in of experimental treatments. The coefficient of determination (R^2) and P-values from this analysis are presented in table 7.13. The linear regression analysis did not find any significant (at the 5% level) correlations between the gene expression and SE-HPLC data for any of the storage protein synthesis gene transcripts analysed.

7.4. **Discussion**

7.4.1. Controlled-environment experiment

The analysis of gluten storage protein synthesis gene expression revealed that the relative expression of these genes is affected by elevated temperature during grain-filling, and that this affect changes as grain-filling progresses. When analysed against the sampling timepoint (or accumulated thermal time after anthesis), five of the six transcripts measured an increase

in expression during mid grain-filling due to elevated temperature, and all six showed a decrease in expression by the end of grain-filling. This suggests a common effect of increased temperature on all of the storage protein synthesis genes investigated as part of the controlled-environment experiment, whereby increased temperature results in a greater peak in expression during mid grain-filling followed by a sharper decrease in expression at the end of grain-filling, confirming the findings of Altenbach *et al.* (2002).

Whilst analysing gluten storage protein synthesis gene expression by accumulated thermal time, and therefore at comparable stages of development, showed that temperature increased expression, when the data was analysed against calendar days, no such increase was found. Comparison of the figures in 7.1 shows that when compared on a day-by-day basis as opposed to by accumulated thermal time, the expression patterns for the two temperature treatments appear overlaid, with little difference between the two treatments. This observation is confirmed by the REML analysis of the data, which failed to identify any significant difference between the two temperature treatments during early grain-filling for any of the six storage protein synthesis gene transcripts analysed. However, whilst there was a reduction in the difference between the temperature treatments during early grain-filling, analysing the expression data against calendar days revealed a greater increase in the effect of elevated temperature at mid to late grain-filling. Although this observation represents a relatively large statistically significant difference between the two treatments, it is perhaps not a particularly valid nor useful observation to make. Whilst the predictions presented in tables 7.1–7.6 report dramatically increasing differences between the two temperature treatments up to 35DPA, it must be remembered that past 25DPA these predictions are projections that reach beyond the time period from which samples were collected from plants grown under the elevated temperature treatment. In fact, at 35DPA plants subjected to the elevated temperature treatment were at full harvest maturity.

In light of previous work (Wan *et al.* 2014), it was expected that there would have been a significant effect of nitrogen input on gluten storage protein synthesis gene expression. However, the mRNA expression results from the controlled-environment experiment fail to identify a significant response to nitrogen input to any of the six storage protein synthesis transcripts analysed. This is perhaps surprising, since the nitrogen treatments used in the controlled-environment resulted in significant differences in protein content and composition (chapter 6), grain protein distribution (chapter 5), and grain yield and associated components (chapter 4). Furthermore, since the elevated temperature treatment had a significant effect on all protein synthesis gene transcripts analysed, the lack of a significant effect from the nitrogen treatments cannot be attributed to excessive variation in the data. Therefore it is likely that this negative result is a genuine reflection of the minimal difference between the high and low nitrogen treatments used (discussed further in section 4.4.1).

Since the analysis of mRNA expression data identified a positive effect on the expression

of storage protein synthesis genes, it is useful to compare this data to the protein content and composition data (presented in chapter 6) to identify and concurrent responses. Across all of the six protein synthesis gene transcripts analysed, a similar response to elevated temperature was apparent: an earlier peak in expression (relative to accumulated thermal time), followed by sharper drop in expression. The effect of this increase in gene expression at mid grain-filling can be observed in the grain protein content results (figure 6.1), whereby grain protein content increases from mid grain-filling onwards under the elevated temperature treatment only. Although in this dataset the drop in expression observed in the elevated temperature treatment is not represented, it may be that by this point in grain-filling this drop in expression is inconsequential, as the majority of protein present at maturity has already been produced. Whilst it is possible to make such casual observations, it must be considered that protein content is a product of both the amount of protein present and the dilution of that protein with other cellular components, and therefore is not the best measure of the physical amount of protein present within the grain. However, the SE-HPLC analysis completed on the mature grain from this study does provide an absolute measurement of protein, and furthermore provides quantification of the different protein groups present. In the linear regression analysis of gene expression and SE-HPLC data (presented in table 7.7), clear correlations where identified between the level of gene expression and the amount of the relevant storage protein in the mature grain. This adds confidence to the findings of both the mRNA expression and SE-HPLC analyses completed as part of the controlled-environment experiment, since it shows that the amount of protein present in the mature grain is a product of the expression of the relevant synthesis genes.

7.4.2. WGIN diversity field trial experiment

Analysis of the mRNA expression data from the WGIN diversity field trial experiment identified the effect of season-to-season variation and different nitrogen input levels on wheat grain storage protein synthesis genes in Cadenza. In contrast to the controlled-environment experiment, the results were comparable whether analysed against accumulated thermal time or sampling timepoint. This is to be expected, since the difference in average temperature between the two years of the field experiment was small in comparison to the difference between the temperature treatments applied in the controlled-environment experiment. As a result there was a minimal difference between years in the accumulated thermal time at each sampling timepoint. Therefore, only the analyses of gene expression in relation to sampling timepoint (calendar days after anthesis) will be discussed, since these results are based on actual datapoints with no predictions or extrapolation beyond the bounds of measurement.

The effect of year on the expression of storage protein synthesis genes was generally a reduction in expression early in grain-filling in 2017 when compared to 2016. This effect was observed in all of the gene transcripts analysed with the exception of γ-gliadin (1), from which the SEM of the data was particularly high, signifying an excessively variable dataset. In the case of the

α- and β-gliadins, γ-gliadin (2), and HMW-glutenin transcripts, a non-linear response was identified over time, with gene expression in 2017 lower at the start and end of grain-filling, but comparable at mid grain-filling (T3–T4). Again, as in the effect of the elevated-temperature treatment in the controlled-environment experiment, the effect of year in the field experiment is conserved across all gene transcripts, suggesting a shared response.

Of the six storage protein synthesis genes analysed, the transcripts for α- and β-gliadins, HMW-glutenins, and ω-gliadins showed a significant response to an increase in nitrogen input (from 100kg-N/ha to 350kg-N/ha). However, when the mean NRQ values were compared with the LSD at the 5% level, only the expression of the ω-gliadin transcript was significantly different between the same sampling timepoint. In this instance, the high nitrogen treatment resulted in an increase in the relative gene expression, but only at the end of grain-filling (T4–T5). Since these comparisons between gene expression measurements at the same sampling timepoint are the most relevant comparisons to make, only the results of the ω-gliadin transcript should be interpreted with a degree of confidence. As in the results from the controlled-environment experiment, this lack of a strong effect of nitrogen input on gene expression is unexpected. However, unlike in the controlled-environment experiment, the lack of a significant response to nitrogen input unlikely to be due to the treatments applied, since the difference between 100kg-N/ha and 350kg-N/ha is considerable, resulting in significant differences in grain yield (figure 4.7a), nitrogen content (figure 6.12b), and protein composition (figure 6.15 and 6.18). Rather, it is likely that these results are due the variability inherent in data collected from field-grown plants.

Comparison of the raw gene expression data presented in figure 7.2 with the data from the controlled-environment experiment (figure 7.1) shows some dramatic differences. Whilst all six gene transcripts show similar expression patterns in the controlled-environment experiment (an early peak in expression that rapidly drops off as grain-filling progresses), there are no such patterns present in the data from the field experiment. Again, this is likely a product of the variability of field-grown samples, and may go some way to explaining the weak results from this experiment.

Perhaps most detrimental to the reliability of the conclusions made from the gene expression results from the WGIN diversity field trial experiment is in the results of the linear regression analysis between protein synthesis gene expression and the SE-HPLC data (see table 7.13). Whilst in the controlled-environment experiment the results from this analysis identified a close relationship between the abundance of each protein in mature grain with the expression of the equivalent synthesis gene, no such correlation was found for any of the gene transcripts analysed in the WGIN diversity experiment (see table 7.13). However, it should be noted that these negative results are a combination of the quality of both the SE-HPLC and the mRNA expression analysis data.

Chapter 8: General discussion

Climate change is perhaps the greatest challenge to ever face mankind, and to date little action has been made to counter the negative effects it will have on our planet. The danger posed by climate change is not only due to the general increase in global temperature, but also due to the predicted increase in the frequency of extreme weather events such as droughts and heatwaves (Hennessy *et al.* 2008). The effect of these localised events will be detrimental to agriculture (Parry *et al.* 2004; Lobell *et al.* 2007; Wheeler *et al.* 2013; Rosenzweig *et al.* 2014), and in the context of wheat production will result in decreased yield (Altenbach *et al.* 2003; Gooding *et al.* 2003; Shah *et al.* 2003) and changes to grain quality and protein content (Gooding *et al.* 2003; Dupont *et al.* 2006b; Dupont *et al.* 2006a; Yang *et al.* 2011). When the issue of climate change is combined with a rapidly growing population, a perfect storm is created whereby more food must be produced in increasingly hostile conditions. This issue presents an even greater challenge when considering the likelihood that regulatory efforts to minimise the environmental impact of agriculture will limit both the chemical inputs on which crop production is so heavily dependant and also the amount of new land that can be converted to agricultural use. Therefore, in the near future crop researchers and breeders must produce crops which tolerate extreme climactic conditions, and are able to produce higher yields under lower inputs of fertilisers and pesticides, without sacrificing quality.

The edible product of the wheat crop is its grain, a single-seeded fruit which is rich in protein and carbohydrates and contains a large starchy endosperm encased within the aleurone and bran layers. During milling for the production of white flour, the endosperm tissue is separated from the aleurone and bran layers, and ground to a fine powder. In the UK, the most popular wheat product is bread produced from white flour, and so understanding how the composition of the wheat endosperm is likely to be effected by climate change is of great importance. To this aim, this PhD thesis analysed wheat grain grown both in controlled-environment and in the field to identify the effects of climate, temperature, nitrogen input, and genotype on the distribution of storage protein within the endosperm, the protein composition and related gene expression, and on the yield and yield components of the wheat crop.

The distribution of protein within the wheat endosperm isn't homogeneous, and it has long been observed that there is a gradient in protein, with a higher concentration of protein towards the outside of the grain (Cobb 1905). This distribution gradient results in the production of mill streams with different protein content, and therefore backing quality (Wang *et al.* 2007), and, since the extraction rate of endosperm from aleurone and bran layers is never 100%, also results in a disproportional amount of protein being removed with the aleurone layer during milling. Therefore any factors that result in an increase in this protein distribution gradient, *i.e.* more protein positioned closer to the aleurone layer, would result in a decrease in the amount of protein recovered in the production of white flour. Since previous studies

on the distribution of protein within the wheat endosperm use low-throughput techniques, one of the aims of this study was to develop a high-throughput method for the quantification of protein distribution gradients in microscopy images. Once developed, this method (Savill *et al.* 2018) was used to generate data on both the gradient in total protein concentration, and the size-distribution of individual protein bodies within the endosperm. In addition to the analysis of protein distribution in the endosperm, measurements were taken for grain yield and yield components, protein content and composition, and for storage protein synthesis gene expression.

This thesis reports the results obtained from plant material from two experiments, a controlled-environment study, and from three years of the ongoing WGIN diversity field trial experiment. The controlled-environment experiment was planned to recreate the effect of a prolonged heatwave during grain-filling, and to identify any interactions that may exist with the supply of nitrogen to the plants prior to anthesis. Whilst the controlled-environment experiment included a single wheat variety (Cadenza), grain was sampled from four varieties from the WGIN diversity field experiment. The aim of the WGIN field experiment was to identify the effect of year-to-year climactic variation (primarily with regards to average temperature during grain-filling) and nitrogen fertiliser regime on a range of wheat genotypes, with high-protein bread-making wheat varieties Hereward and Cadenza, early-flowering Soissons, and low-protein feed wheat Istabraq all sampled over three years of field experiments. This general discussion chapter discusses key findings (section 8.1), limitations of the study (section 8.2), proposes topics for further investigation (section 8.3), and provides a conclusion of the study (section 8.4), as well as some more general concluding remarks (section 8.5).

8.1. Key findings

Elevated temperature during grain-filling and increased nitrogen supply increase the gradient in protein concentration across the wheat endosperm. A novel finding of the controlled-environment experiment was that increased temperature during grain-filling results in an increase in the gradient of protein in the wheat grain, with more protein concentrated in the outer layers of the endosperm (Savill *et al.* 2018). Furthermore, an interaction with the amount of nitrogen supplied during vegetative growth was identified, whereby high nitrogen input resulted in a minimal increase in the protein distribution gradient under control (20°C) temperatures, but a considerable increase under elevated (28°C) temperatures. A similar effect of nitrogen input on the grain protein distribution gradient was also identified in the WGIN diversity field experiment, with greater gradients in grain grown under the high (350kg-N/ha) nitrogen treatment. As part of the field experiment, samples from 2015 and 2017 were analysed for grain protein distribution gradients, and although the difference between these two years was minimal, there was generally a slight increase in the response to nitrogen in 2017, the warmer of the two years. In summary, these findings predict the effect that a prolonged heatwave during grain-filling is likely to have on the protein distribution within the wheat grain endosperm: the increased temperature will result in more protein concentrated

in the endosperm closest to the aleurone layer, but decreasing the input of nitrogen prior to anthesis may negate this effect to a certain degree. Although predicting a heatwave during grain-filling is impossible at the point of nitrogen application, this information may be useful in years where there is a higher chance of increased summer temperatures, for example in coincidence with El Niño/the North Atlantic Oscillation weather phenomena, which has already been linked to the performance of wheat crops in the UK (Kettlewell *et al.* 1999).

The size-distribution of protein bodies in the wheat endosperm is affected by nitrogen supply. The gradient in total protein concentration across the wheat endosperm is accompanied by a gradient in the average size of protein bodies. Whilst previous work has identified a link between the differences in average protein body size in the outer and inner endosperm due to drought (Chen *et al.* 2016), the present study is the first to show a link with nitrogen fertiliser input and protein body size-distribution (Savill *et al.* 2018). These results, from both controlled-environment and field experiments, demonstrate that increasing nitrogen supply results in an increase in the gradient in protein body size-distribution (see figures 5.2 and 5.3). Although the effects of post-anthesis temperature and year-to-year climatic variations were also investigated, the effect of these factors was either largely inconsistent, or too small to provide robust conclusions. The effect of elevated temperature on protein body size-distribution in the controlled-environment experiment was particularly inconsistent, with the effect of temperature apparently reversing between the two sampling timepoints. Therefore it is impossible to accurately summarise the effect of elevated temperature on protein body size-distribution. However, the effect of nitrogen input is consistent across both controlled-environment and field experiments, and by analysing the histograms representing the protein body size-distribution data it is apparent that the effect of increased nitrogen supply is an increase in the relative abundance of the smallest protein bodies within the central endosperm. These small protein bodies are likely newly formed, and an increase in their abundance could signify an increase in the initiation of new protein bodies. However, since the protein body size-distribution analysis is unable to generate reliable absolute quantitative data, it is not possible to specify the causes of these changes in size-distribution with any degree of confidence, *i.e.* the increase in the relative abundance of the smallest protein bodies in the inner endosperm could be due to either an increase in the number of small protein bodies, or due to a decrease in the number of larger protein bodies. Regardless, the significance of these results is that they provide further insight into the process of protein accumulation in the wheat endosperm, and how this process is affected by the supply of nitrogen to the plant.

Grain storage protein accumulation varies between genotypes. Whilst the total protein concentration and protein body size-distribution gradient analyses identified general effects of elevated temperature and nitrogen input, there were also differences observed between the four varieties sampled as part of the WGIN diversity field trial experiment. These four varieties were the NABIM group one bread-making wheat Hereward, group two bread-making wheats Cadenza and Soissons, and the low-protein group four feed wheat Istabraq.

With regards to the gradient in total protein concentration, Istabraq showed the least difference in protein concentration across the endosperm, and also the smallest response to increased nitrogen fertiliser input. Conversely, the high-protein varieties Cadenza and Hereward recorded the largest protein concentration gradients, as well as the greatest response to nitrogen input. The differential responses to nitrogen input showed by each genotype were largely replicated in the protein body size-distribution results, with the largest response to increased nitrogen input recorded in Cadenza, and with Istabraq failing to show any response to nitrogen input. These results suggest a relationship between total grain protein content, and the gradient in the distribution of this protein within the grain.

Grain protein content is increased by nitrogen fertilisation and high temperatures during grain-filling. The protein content (as percentage dry matter) of mature wheat grain was increased by both increased temperatures during grain-filling and by increased nitrogen supply prior to anthesis in the controlled-environment, with an interaction between temperature and nitrogen supply whereby increased nitrogen input had a greater effect on grain protein content when temperatures were increased during grain-filling. In the field experiment grain protein content was higher in warmer years, as well as in plants provided with the high (350kg-N/ha) nitrogen treatment, however no interaction between year of experiment and nitrogen treatment was identified. This effect of temperature and nitrogen supply on wheat grain protein content is already well know, with the result of this study adding to the findings of numerous other studies (Dupont *et al.* 2006b; Kindred *et al.* 2008; Nakano *et al.* 2008).

Year-to-year variation in grain protein content is determined by genotype. Whilst there effect of increasing nitrogen supply on the protein content of the mature grain was similar in each of the four varieties sampled in the WGIN diversity field experiment, these effects were more consistent between years for certain genotypes. Consistency between years is a particularly useful trait, and is often marks the difference between NABIM group one and group two bread-making wheats, whereby group one wheats are most consistent in their performance. Therefore such consistency is importance, since it stabilises the income of farmers and provides a consistent product from which millers can produce flour. Of the four genotypes examined in the field experiment, Istabraq was the most consistent in terms of grain protein content at maturity, closely followed by Cadenza, whilst the grain protein content of Hereward was unexpectedly low in 2016. This result is even more intriguing since the grain yield of Hereward in 2016 was the highest of all four genotypes. Therefore, Hereward may be unique in its performance in 2016, a year of high rainfall, moderate temperatures, and low sunlight, an effect that may be worthy of further study. Whilst Hereward would be expected to be consistent between years as a NABIM group one bread-making wheat, it should be noted that Hereward was removed from the HGCA (now AHDB) list of recommended winter wheat varieties in 2011.

Grain protein composition is affected by nitrogen input and post-anthesis temperature. The protein composition (as measured by SDS-PAGE and SE-HPLC) results generated in this study were generally inconsistent, with either a lack of significant differences between treatments, or contradictions between the two methods of analysis or between the two experiments. The effect of the elevated temperature treatment in the controlled-environment experiment only identified one significant results from both SDS-PAGE and SE-HPLC analyses: increased temperature during grain-filling increases the relative abundance of HMW-glutenin proteins, which confirms the findings of Hurkman *et al.* (2013). With regards to the effect of nitrogen, there was a greater degree of correlation between the controlled-environment and field experiment results. In both experiments, increased nitrogen input resulted in an increase in the proportion of α-, β-, and γ-gliadins, ω-gliadins, but a decrease in the proportion of LMW-glutenins, again supporting the findings of Hurkman *et al.* (2013) and Wan *et al.* (2014). Between the years of the WGIN diversity field trial experiment, the only significant effect was on the relative abundance of ω-gliadins as detected by SE-HPLC, which was lowest in 2015, and highest in 2017. However the significance of this effect was borderline (P=0.040). Whilst these results provide some insight to the effects of elevated temperature during anthesis, year-to-year variation in the field, and nitrogen supply on the protein composition of the wheat grain, it does not provide a complete picture of how these factors impact on the accumulation of individual proteins. Due to the numerous contradictions and inconsistencies in theese protein composition results, the analysis of grain gluten content and gluten quality derived from the comparison of SE-HPLC fraction data are perhaps more relevant for predicting the bread-making quality of grain sampled in this study.

Wheat bread-making quality is improved by high temperatures and increased nitrogen input. In terms of grain protein, the determining factors of bread-making quality are gluten content, and the quality of that gluten (El Haddad *et al.* 1995; Sapirstein *et al.* 1998). Under both field and controlled-environment conditions, gluten content was increased at higher levels of nitrogen input, confirming previous reports (Malik *et al.* 2011; Moldestad *et al.* 2014). Additionally, in the controlled-environment experiment, elevated temperature was found to increase grain gluten content, again confirming previous work (Malik *et al.* 2011; Moldestad *et al.* 2014; Tao *et al.* 2018). This increase in gluten content was accompanied by an increase in gluten quality as determined by the comparison of SE-HPLC fractions F1/F2 (the ratio between HMW- and LMW-glutenins) and (F3+F4)/F1 (the ratio between gliadins and HMW-glutenins), presented in figure 6.9, a comparison that can be used as a predictor of bread-making quality (Millar 2003). These results indicate that bread-making quality is improved when temperatures are increased during grain-filling, and also that increasing nitrogen input greatly improves bread-making quality under control temperature conditions, but only slightly improves it when temperatures are higher. These results suggest that nitrogen input is less of a determinant of bread-making quality under high temperatures, and show that

lowering nitrogen input to reduce the gradient in grain protein distribution when temperatures are high during grain-filling would not be to the detriment of bread-making quality.

Expression patterns of gluten storage protein synthesis genes in Cadenza are altered by elevated temperature during grain-filling under controlled-environment conditions. For the six gluten storage protein (α- and β-gliadins, γ-gliadins (1 and 2), HMW-glutenins, LMW-glutenins, and ω-gliadins) synthesis gene transcripts analysed in spring wheat Cadenza, a common response to elevated temperature was identified: gene expression reaches a higher, earlier peak, before dropping off more rapidly towards the end of grain-filling, a response previously described by (Altenbach *et al.* 2002). The gene expression from the controlled-environment was analysed both in terms of sampling timepoint (equivalent to accumulated thermal time) and in terms of calendar days. With the adjustments for thermal time removed from the analysis, the effect of temperature treatment largely disappear, with the exception of predictions that reach beyond the limits at which measurements were taken from plants grown under the high temperature treatment (*i.e.* predictions after 25DPA). Therefore the analysis of gene expression against sampling timepoint is perhaps the most relevant interpretation of the data, and shows how increased temperature increases the expression of genes involved in gluten protein synthesis in early to mid grain-filling, before reducing it at the final sampling timepoint of grain-filling. To add confidence to the validity of these results, the linear regression of accumulated gene expression data with the absolute measurements of the relevant SE-HPLC fractions showed strong correlations between all of the protein synthesis genes analysed and the amount of the respective protein present in the grain. No significant difference in storage protein synthesis gene expression was identified between the two nitrogen treatments applied, a results that is likely due to the minimal difference between the low and high nitrogen treatments used in this experiment.

Gluten synthesis gene expression in of field-grown Cadenza varies from year-to-year, and is affected by nitrogen input. In 2017, the expression of gluten storage protein synthesis gene transcripts during early to mid grain-filling was lower than in 2016. However, the effect of nitrogen was less pronounced, and although significant interactions were identified between nitrogen treatment and the expression of α- and β- gliadin, HMW-glutenin, and ω-gliadin transcripts, direct comparisons between sampling timepoints revealed that the only significant differences at the 5% level were from the expression of ω-gliadins, which was increased at the later stages of grain-filling, supporting the previous findings of Wan *et al.* (2014). Comparison of the expression patterns from the field-grown material (presented in figure 7.2) with the somewhat uniform expression patterns from the controlled-environment experiment (figure 7.1) give some indication of the reason behind the lack of significant results from the mRNA expression analysis of the WGIN diversity field trial grain. The expression patterns from the field-grown wheat are erratic, an observation supported by the linear regression analysis between the gene expression and SE-HPLC protein data, which fails to find any significant correlations between the two datasets. The most likely explanation

between the unreliability of these results is the variation which is inherent in field-grown plants, a limitation that is always difficult to overcome in field-based experiments.

Grain yield is increased by nitrogen fertilisation, and reduced by higher temperatures during grain-filling. In both the field and the controlled-environment experiments, yield was increased by providing more nitrogen to the plants, and was reduced under both the elevated temperature treatment in the controlled-environment experiment and in the warmer years of the field trial, effects which have been reported many times before (Thorne *et al.* 1987; Mitchell *et al.* 1993; Kindred *et al.* 2008). To explain the effect of temperature and nitrogen on grain yield in the controlled-environment experiment further, TGW was measured. In the analysis of data from the controlled-environment experiment, it was determined that the reduction in grain yield due to the elevated temperature treatment could be principally attributed to a reduction in grain size (TGW). However, it was also apparent that whilst limiting nitrogen supply reduced TGW, it also reduced number of grain produced, and as such the reduction in grain yield under the low nitrogen treatment was a product of fewer, smaller grain being produced.

Limiting nitrogen supply produces shorter grains, and elevated temperature during grain-filling produces narrower grains. Grain from the controlled-environment experiment was analysed for area, length, and width to identify the effect of elevated temperature during grain-filling and nitrogen supply on grain morphology. Whilst the measurements for grain area approximated the TGW results, the grain length and width measurements showed differential responses to the temperature and nitrogen treatments: grain length increased 3% under the high nitrogen treatment (contary to the findings of Kindred *et al.* (2008)), and grain width was reduced by 12% by the elevated temperature treatment. The difference in grain width due to increased temperature is particularly relevant, since the increase in protein distribution gradients observed in grain subjected to the elevated temperature treatment are potentially a product of the reduced width of these grain, which effectively reduces the distance over which the protein distribution gradient is established, resulting in a steeper gradient.

Early-flowering Soissons achieved lower yields and produced smaller grain in 2016 than 2015. French bread-making wheat Soissons was included as part of the WGIN diversity field trial experiment since as a photoperiod-insensitive variety it reaches anthesis, and therefore begins grain-filling, considerably earlier than the other varieties used in the experiment (see table 3.1). As a result, in 2016 Soissons accumulated less thermal time than Cadenza, whilst in 2015 a comparable amount of thermal time was accumulated over the first 35 days of grain-filling (see figure 3.5). Soissons was included in this study in an effort to identify any effects that this early-flowering phenotype might have both during grain-filling and at maturity. However due to the lack of any form of control, it is impossible to attribute any differences between the performance of Soissons and other wheat genotypes to

its early-flowering phenotype alone. Therefore only a casual summary of potentially relevant observations are presented. Additionally, since no samples were available from Soissons harvested at maturity in 2017, only the differences between 2015 and 2016 can be made. With regards to the differences in accumulated thermal time for Soissons and the other genotypes in 2016 compared to 2015, the most relevant observation is the fact that Soissons achieved the lowest in 2016, compared to the second lower grain yields in 2015. Additionally, Soissons was the only variety that recorded lower TGW in 2016 compared to 2015, producing much smaller grain in 2016 than the other varieties. In summary, whilst it is impossible to assign a causation, Soissons did appear to suffer in 2016, producing smaller grain, and ultimately achieving a lower yield.

8.2. **Limitations of the study**

Whilst this thesis presents novel findings on the factors that affect the distribution of protein in the wheat endosperm, there are limitations both in these findings and in the results of the protein composition and gene expression analyses presented. These limitations include compromises made in both the design of experiments, and in the processing and analysis of samples.

The biggest compromise made in the design of the controlled-environment experiment is the lack of replication of the whole experiment. Since only two controlled-environment rooms were used for the experiment, one for the control and one for the elevated temperature treatment, it could be argued that any effects observed in plants subjected to the elevated temperature treatment are as a result of the different room used rather than any treatment applied. To remove any room-effect, the entire experiment should have been repeated with the controlled-environment rooms switched. However this approach would be prohibitively expensive, and since the two rooms used in this experiment were both high-quality modern controlled-environment rooms of identical specification with quality assurance protocols in place to monitor lighting levels, the author is confident that any difference observed between the two rooms can be accounted for by the difference in temperature alone. A limitation that impacted on the results of the experiment, and was not apparent at the time the experiment was run, was the limited difference between the two nitrogen treatments used. "Rothamsted nematode mix" potting mix was used in this experiment, with a full and one-tenth strength liquid nitrogen fertiliser applied periodically during vegetative development. The decision to use this potting mix was informed by the work of Derkx *et al.* (2012), who used the same nutrient solution composition in combination with the Rothamsted nematode potting mix to achieve significant differences between wheat plants provided with the high- and low-nitrogen treatments, with plants subjected to the low-nitrogen treatment showing reduced tillering. In our experiment there was no visual difference between the plants supplied with the low nitrogen treatment and the plants supplied with the high nitrogen treatment. However, the leaf chlorophyl content analysis completed on plants at the time of anthesis did show a minimal, yet statistically significant, difference in the chlorophyll content of the plants of each

nitrogen treatment (see in section 4.2.5). After consultation with the potting mix supplier (Petersfield Products, Leicester, UK), it is apparent that the loam component of the mix had a full-spectrum fertiliser applied, an addition that was likely omitted upon the request of the authors of the Derkx *et al.* (2012) study. As a result, the potting mix was had a greater nutrient content than anticipated, which effectively rendered the low- and high-nitrogen treatments as "sufficient" and "abundant" rather than 'deficient' and 'sufficient'.

The WGIN diversity field trial was used to provide grain samples from four wheat varieties over three years of field trials (2015–2017), in an effort to identify how the response to nitrogen fertiliser varies between difference genotypes over multiple years. However, due to the variable nature of data collected from field-based experiments it was often difficult to draw conclusions from the data. This variability was also likely exacerbated by the fact that the field trial was hosted on a different site each year. Regardless of any issues surrounding the variability of samples collected, the purpose of running a multi-year field experiment was to identify the effect of temperature variation between years on the process of grain-filling. However, any differences between years cannot be attributed solely to differences in temperature as in the controlled-environment experiment, but rather could be due to any number of factors, including other climatic conditions, the prevalence of pests, lodging, and variation in the nutritional composition of the soil in each field. Therefore any comparisons of the results between years can only be made on a speculative basis, and never attributed to a single factor such as temperature during grain-filling. However, whenever supporting evidence is available from the controlled-environment experiment, the validity of such comparisons is increased. A final limitation relating to the used of the WGIN diversity field trial experiment during this study relates to the scope of the sampling completed. The WGIN diversity field trial has run since 2004, and included 20–30 commercial wheat varieties grown under one of four different levels of applied nitrogen fertiliser. Due to time constraints, only four of the available genotypes were sampled, under two of the four nitrogen treatments. Analysing more genotypes grown under more levels of nitrogen fertiliser would have added to the results, and may have provided some interesting exceptions to the results presented in this thesis.

A limitation in the sampling method employed in both the controlled-environment and the WGIN field trial experiment was the frequency with which samples were collected for microscopy analysis. Whilst gene expression and protein content data was easily generated from all five of the sampling timepoints used, this was not possible with the light-microscopy analysis. As a result, light-microscopy images were only analysed at two timepoints in the controlled-environment experiment, and at a single timepoint in the field experiment. This is due to the narrow time-frame from which developing grain can be processed and sectioned. Grain sampled too late will not section cleanly, and grain sampled too early will contain insufficient protein to analyse. In hindsight, grain should have been sampled more frequently than every seven days for microscopy analysis, allowing for a true timecourse analysis of grain protein accumulation over time. Likewise, the loss of a years worth of field experiment samples

due to improper processing reduced the size of the potential dataset by a third, decreasing the changes of identifying a year effect on grain protein distribution.

Although the combination of image analysis and light-microscopy produced a large, reliable dataset which characterised both the gradients in protein concentration and size-distribution of protein bodies in the wheat endosperm, due to the limitations imposed by the processing of plant tissue for light microscopy analysis, only developing grain could be analysed. Whilst the grain processed for light-microscopy analysis was sampled as late in grain-filling as possible to produce good quality microscopy images, the images which were analysed were not of mature grains, and the lack of grain moisture content data mean that the exact developmental maturity of the sampled grain is uncertain. Therefore the observations made on the protein distribution gradients in these grain must be extrapolated to made assumptions on mature grain. Although this is an obvious concern, the results of the grain concentration gradient analysis from the controlled-environment experiment was completed on grain sampled at two timepoints during development, and the effects observed increase between the early and later timepoint, suggesting that the observations are likely to be preserved until the end of grain-filling. This observation, combined with previous studies demonstrating the presence of a protein distribution gradient in mature grain (Cobb 1905; Kent 1966; He *et al.* 2013), provide a certain level of confidence that the observations made at mid to late grain-filling as part of this study are applicable to mature grain.

A consideration to be made when interpreting the protein body size-distribution results in that the sections imaged and analysed were collected a considerable distance apart (at least 20 µm), in order to sample entirely different protein bodies in each section. Due to this approach, there is no three-dimensional information available. Therefore the smallest protein bodies detected could represent newly-formed protein bodies, or they could equally represent the extreme of a much larger protein body. This potential source of measurement error is present in protein bodies of all sizes, since with a 1 µm section, it is unlikely that the maximum width of a roughly-spherical protein body has been captured. Indeed it is most likely that the protein body area measures does not represent the maximum size of that protein body. Whilst sub-ideal, the sectioning methodology used was borne out of necessity, since collecting serial whole-grain sections of sufficient quality is practically impossible.

With regards to the methods of analysis chosen for the grain samples collected, analysing protein composition and associated gene expression in whole-grain flour samples isn't relevant to the distribution of these proteins within the endosperm. Whilst the protein composition and gene expression data presented in this thesis is an interesting addition to the microscopy analysis of grain protein distribution, it did not investigate the nature of protein distribution in the wheat grain endosperm. Suitable alternatives to the protein composition and gene expression analysis approach taken in this study would have been the application of the image-analysis software used in this study on immunofluorescence-microscopy and *in situ*

hybridisation images respectively to investigate the spatial distribution of specific proteins, and related gene expression (as discussed in section 8.3). Such an approach would allow for the quantification of gradients in different storage proteins, and also the spatial analysis of gene expression related to the synthesis of these proteins.

8.3. **Future work**

The greatest scope for work continuing on from this study is in the further application of the image analysis software method that was developed for the quantification of protein distribution gradients (Savill *et al.* 2018). Whilst this method was used to measure gradients in both total protein concentration and protein body size-distribution, it can be applied to any images in which there is clear contrast between areas of interest and background. Without any adaptation, this image analysis technique could be used to quantify (either spatially or generally) any number of cellular components within microscopy images. In the context of furthering the finding presented in this thesis, however, the most relevant applications would be in the spatial analysis of immunofluorescence-microscopy and *in situ* hybridisation images of wheat endosperm tissue.

Since different wheat storage proteins are known to show different distribution patterns across the endosperm (Wang *et al.* 2007; Tosi *et al.* 2011; He *et al.* 2013; Wan *et al.* 2014), an obvious continuation of the present study would be to use immunofluorescence microscopy to investigate how climate, temperature, and nitrogen input affect the distribution of the different storage proteins. Such work has already been completed on ω-gliadins, with Wan *et al.* (2014) showing that increased nitrogen input results in an increase in the accumulation of ω-gliadins in the outer endosperm. The high-throughput analysis technique presented in this study was tested on immunofluorescence-microscopy images during its development, and is a suitable means to generate data on the distribution of individual storage proteins identified through immunofluorescence microscopy.

Another application of the image analysis technique developed would be in the spatial analysis of gene expression through the analysis of *in situ* hybridisation images. Such an approach could be used in experiments similar to those conducted by Drea *et al.* (2005) and Wan *et al.* (2014), and could facilitate much larger studies on the localisation of gene expression within the wheat endosperm. As an extension to the present study, *in situ* hybridisation could be used to link gene expression data with protein distribution data, something that was not possible with the approach taken here. When completed in combination with immunofluorescence microscopy, such a study could provide a complete picture of the synthesis and accumulation of different gluten storage proteins within the wheat endosperm.

In terms of furthering the present study, and adding to the dataset of grain protein gradients measured from light-microscopy image, there is definite value in continuing to analyse the gradients in grain protein from subsequent years of the WGIN diversity field trial. Adding data from additional years would facilitate the application of multivariate analysis, which

could statistically identify the effect of individual climatic factors, including hours of sunlight, temperature, and rainfall. Furthermore, collecting data from more years increases the chances of capturing data during an extreme weather event, such as the record-breaking heatwave experienced in the UK in 2018. In addition to continuing this experiment over future years, the number of genotypes sampled could be increased, and samples could be collected from all four nitrogen treatments (0kg-N/ha, 100kg-N/ha, 200kg-N/ha, and 350kg-N/ha). The expansion and continuation of this experiment would certainly add to the results already produced, and might aid in identifying genotypes that are more resistance to the increased temperatures expected to be commonplace in the future.

A more practical analysis of the impact of differences in the distribution of protein in the endosperm with regards to the quality of the mature grain would be to employ experimental milling techniques. Whilst pearl-milling was used by He *et al.* (2013) to identify a response to nitrogen input in the distribution of both total protein and individual protein groups in the wheat grain, this technique isn't comparable to commercial milling. The use of an experimental roller mill (a miniature version of a commercial mill), as used by Wang *et al.* (2007), produces multiple mill streams which are more analogous to those produced in a commercial mill, with true bran-separation, and the production of flour enriched with different parts of the grain. Analysing the effect that nitrogen input and temperature during grain-filling have on the behaviour of grain in an experimental mill would provide more applicable information to millers and bakers, and is another logical succession to the information presented in this thesis. However, experimental milling requires large grain samples, which whilst not an issue for field-based experiments, could render controlled-environment experiments prohibitively expensive. As a continuation of experimental milling, test-loaves of bread could be baked to test for bread-making, rather than relying on predictions based on the results of SE-HPLC protein composition analysis which only consider the role of protein composition as a determinant of bread-making quality.

With regards to the effects that climate change will have on the quality of wheat grain in the future, the present study focusses solely on an increase in temperature during grain-filling. However, climate change will also bring about an increase in the likelihood of drought, and is closely associated with increased levels of atmospheric CO_2. Both of these factors are also know to have an effect on wheat, with drought reducing yield (Shah *et al.* 2003) and altering protein content (Altenbach *et al.* 2003; Yang *et al.* 2011), and increased CO_2 linked to an increase in yield (largely due to an increase in starch accumulation) and a reduction in bread-making quality (Fernando *et al.* 2015). Since there are known interactions between elevated temperature, drought, and increased levels of atmospheric CO_2 on the physiology of the wheat plant, it would be interesting to investigate how these factors also affect grain protein distribution.

8.4. Conclusion of experiment

Although the effects of climate change will be unprecedented and indiscriminate, thanks to decades of research, they will not take the scientific community by surprise. To support a growing population under increasingly difficult circumstances, agricultural research must continue to focus on producing resilient crops which are able to withstand extreme conditions whilst remaining productive and nutritious. To this aim, the effect of both prolonged heatwave temperatures and restricted nitrogen input on the distribution, content, and composition of protein within the wheat grain was investigated under controlled-environment conditions. Multiple wheat varieties were also grown in the field, and the effect of year-to-year climatic variation and nitrogen fertiliser regime analysed to identify any differential responses between genotypes. A major outcome of this study is a better understanding of how elevated temperature and nitrogen fertilisation interact to increase the protein distribution gradient in the wheat grain endosperm, resulting in a greater proportion of protein accumulated in the outer endosperm. Furthermore, a novel image analysis technique has been developed and made publicly available (Savill *et al.* 2018) which has the scope to greatly increase the throughput of the spatial analysis of objects in microscopy images, facilitating the completion of much larger experiments reliant on these techniques in the future.

8.5. Concluding remarks

The past four years have marked a great political change across the world.

In 2015, 195 countries signed up to The Paris Agreement: the first ever legally-binding agreement for all nations to "undertake ambitious efforts to combat climate change", with a shared aim of limiting the global increase in temperature to 1.5°C.

In 2017, the President of the United States announced his intention to withdraw from The Paris Agreement; climate modelling studies have predicted that global warming will almost certainly exceed 1.5°C (Brown *et al.* 2017; Mauritsen *et al.* 2017; Raftery *et al.* 2017); and it has emerged that the industrialised nations participating in The Paris Agreement are all failing to honour their promises regarding reductions in greenhouse gas emissions (Victor *et al.* 2017).

With the leaders of the world continuing to prioritise their own political and financial ambitions over the future of our planet, and the sweeping rise in anti-intellectualism and right-wing populism in recent years, it is difficult to be optimistic about the future.

I'm not a believer in global warming.

And I'm not a believer in man-made global warming.

It could be warming, and it's going to start to cool at some point.

And you know, in the early [sic], in the 1920s, people talked about global cooling...

They thought the Earth was cooling.

Now, it's global warming...

But the problem we have, and if you look at our energy costs, and all of the things that we're doing to solve a problem that I don't think in any major fashion exists [sic].

Donald Trump – *45th president of the United States*

Bibliography

AHDB (2018). Wheat growth guide. *https://cereals.ahdb.org.uk/media/185687/g66-wheat-growth-guide.pdf*, 1–44.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L. (2006). Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* **111**:

Altenbach, S. B., Kothari, K. M., and Lieu, D. (2002). Environmental conditions during wheat grain development alter temporal regulation of major gluten protein genes. *Cereal Chemistry* **79**: 279–285.

Altenbach, S. B., DuPont, F. M., Kothari, K. M., Chan, R., Johnson, E. L., and Lieu, D. (2003). Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. *Journal of Cereal Science* **37**: 9–20.

Amthor, J. S. (2001). Effects of atmospheric CO2 concentration on wheat yield: Review of results from experiments using various approaches to control CO2 concentration. *Field Crops Research* **73**: 1–34.

Asseng, S., Foster, I., and Turner, N. C. (2011). The impact of temperature variability on wheat yields. *Global Change Biology* **17**: 997–1012.

Asseng, S., Ewert, F., Martre, P., Rötter, R. P., Lobell, D. B., Cammarano, D., Kimball, B. A., Ottman, M. J., Wall, W., White, J. W., Reynolds, M. P., Alderman, P. D., Prasad, P. V. V., Aggarwal, P. K., Anothai, J., Basso, B., Biernath, C., Challinor, A. J., De Sanctis, G., Doltra, J., Fereres, E., Garcia-Vila, M., Gayler, S., Hoogenboom, G., Hunt, L. A., Izaurralde, R. C., Jabloun, M., Jones, C. D., Kersebaum, K. C., Koehler, A. K., Müller, C., Naresh Kumar, S., Nendel, C., O'Leary, G., Olesen, J. E., Palosuo, T., Priesack, E., Eyshi Rezaei, E., Ruane, A. C., Semenov, M. A., Shcherbak, I., Stöckle, C., Stratonovitch, P., Streck, T., Supit, I., Tao, F., Thorburn, P. J., Waha, K., Wang, E., Wallach, D., Wolf, J., Zhao, Z., and Zhu, Y. (2015). Rising temperatures reduce global wheat production. *Nature Climate Change* **5**: 143–147.

Avery, B. W. and Catt, J. A. (1995). The soil at Rothamsted, pp. 45.

Bancal, P. (2009). Decorrelating source and sink determinism of nitrogen remobilization during grain filling in wheat. *Annals of Botany* **103**: 1315–1324.

Bebber, D. P., Ramotowski, M. A. T., and Gurr, S. J. (2013). Crop pests and pathogens move polewards in a warming world. *Nature Climate Change* **3**: 985–988.

Bechtel, D. B., Zayas, I., Kaleikau, L., and Pomeranz, Y. (1990). Size-distribution of wheat starch granules during endosperm development. *Cereal Chemistry* **67**: 59–63.

Békés, F., Kemény, S., and Morell, M. (2006). An integrated approach to predicting end-product quality of wheat. *European Journal of Agronomy* **25**: 155–162.

Bentley, A. R., Horsnell, R., Werner, C. P., Turner, A. S., Rose, G. A., Bedard, C., Howell, P., and Wilhelm, E. P. (2013). Short, natural, and extended photoperiod response in BC2 F4

lines of bread wheat with different *Photoperiod-1* (*Ppd-1*) alleles. *Journal of Experimental Botany* **64**: 1783–1793.

Bergman, C. J., Gualberto, D. G., Campbell, K. G., Sorrells, M. E., and Finney, P. L. (1998). Genotype and environment effects on wheat quality traits in a population derived from a soft by hard cross. *Cereal Chemistry* **75**: 729–737.

Brown, P. T. and Caldeira, K. (2017). Greater future global warming inferred from Earth's recent energy budget. *Nature* **552**: 45–50.

Carceller, J. L. and Aussenac, T. (1999). Accumulation and changes in molecular size distribution of polymeric proteins in developing grains of hexaploid wheats: Role of the desiccation phase. *Australian Journal of Plant Physiology* **26**: 301–310.

Carson, G. R. and Edwards, N. M. (2009). "Criteria of wheat and flour quality". In: *Wheat Chemistry and Technology*. Ed. by K Khan and P. Shewry. 4th Ed. St Paul: AACC, pp. 97–118.

Chen, X., Li, B., Shao, S. S., Wang, L. L., Zhu, X. W., Yang, Y., Wang, W. J., Yu, X. R., and Xiong, F (2016). Accumulation characteristic of protein bodies in different regions of wheat endosperm under drought stress. *Journal of Integrative Agriculture* **15**: 2921–2930.

Chope, G. A., Wan, Y., Penson, S. P., Bhandari, D. G., Powers, S. J., and Shewry, P. R. (2014). Effects of genotype, season, and nitrogen nutrition on gene expression and protein accumulation in wheat grain. *Journal of Agricultural and Food Chemistry* **62**: 4399–4407.

Chung, O. K., Pomeranz, Y., and Finney, K. F. (1978). Wheat flour lipids in breadmaking. *Cereal Chemistry* **55**: 598–618.

Cobb, N. A. (1905). Universal nomenclature of wheat. *Department of Agriculture - New South Wales, Miscellaneous Publication, No.539*, 56–60.

Cornec, M., Popineau, Y., and Lefebvre, J. (1994). Characterisation of gluten subfractions by SE-HPLC and dynamic rheological analysis in shear. *Journal of Cereal Science* **19**: 131–139.

Courtin, C. M. and Delcour, J. A. (2002). Arabinoxylans and endoxylanases in wheat flour bread-making. *Journal of Cereal Science* **35**: 225–243.

D'appolonia, B. and Rayas-Duarte, P. (1995). "Wheat carbohydrates: Structure and functionality". In: *Wheat Production, Properties and Quality*. Ed. by W Bushuk and V. F. Rasper. 1st. Glasgow: Blackie Academic and Professional, pp. 107–127.

Davidson, E. A. (2009). The contribution of manure and fertilizer nitrogen to atmospheric nitrous oxide since 1860. *Nature Geoscience* **2**: 659–662.

Davis, K., Cain, R., Peters, L., Tourneau, D., and McGinnis, J. (1981). Evaluation of the nutrient composition of wheat. II. Proximate analysis, thiamin, riboflavin, niacin and pyridoxine. *Cereal Chemistry* **58**: 116–120.

Davis-Knight, H., Weightman, R. M., Agu, R., Bringhurst, T., and Brosnan, J. (2007). *Project teport No . 467 Maximising bioethanol processing yield of uk wheat: Effects of non starch polysaccharides in grain*. Tech. rep. HGCA.

Derkx, A. P., Orford, S., Griffiths, S., Foulkes, M. J., and Hawkesford, M. J. (2012). Identification of differentially senescing mutants of wheat and impacts on yield, biomass and nitrogen partitioning. *Journal of Integrative Plant Biology* **54**: 555–566.

Drea, S., Leader, D. J., Arnold, B. C., Shaw, P., Dolan, L., and Doonan, J. (2005). Systematic spatial analysis of gene expression during wheat caryopsis development. *The Plant Cell* **17**: 2172–2185.

Dupont, F. M., Hurkman, W. J., Vensel, W. H., Chan, R., Lopez, R., Tanaka, C. K., and Altenbach, S. B. (2006a). Differential accumulation of sulfur-rich and sulfur-poor wheat flour proteins is affected by temperature and mineral nnutrition during grain development. *Journal of Cereal Science* **44**: 101–112.

Dupont, F. M., Hurkman, W. J., Vensel, W. H., Tanaka, C., Kothari, K. M., Chung, O. K., and Altenbach, S. B. (2006b). Protein accumulation and composition in wheat grains: Effects of mineral nutrients and high temperature. *European Journal of Agronomy* **25**: 96–107.

El Haddad, L., Aussenac, T., Fabre, J., and Sarrafi, A. (1995). Relationships between polymeric glutenin and the quality characteristics for seven common wheats (*Triticum aestivum*) grown in the field and greenhouse. *Cereal Chemistry* **72**: 598–601.

Eliasson, A. C. and Larsson, K. (1990). *Cereals in breadmaking: a molecular colloidal approach.* Marcel Dekker.

Evers, A. D. (1973). The size distribution among starch granules in wheat endosperm. *Starch - Stärke* **25**: 303–304.

FAOSTAT (2014). Food and Agriculture Organization of the United Nations. *http://faostat.fao.org.*

Fernando, N., Panozzo, J., Tausz, M., Norton, R., Fitzgerald, G., Khan, A., and Seneweera, S. (2015). Rising CO2 concentration altered wheat grain proteome and flour rheological characteristics. *Food Chemistry* **170**: 448–454.

Fischer, R. A. (1985). Number of kernels in wheat crops and the influence of solar radiation and temperature. *Journal of Agricultural Science Cambridge* **105**: 447–461.

Fisher, D. B. and Macnicol, P. K. (1986). Amino acid composition along the transport pathway during grain filling in wheat. *Plant Physiology* **82**: 1019–1023.

Gan, Z., Ellis, P. R., and Schofield, J. D. (1995). Gas cell stabilisation and gas retention in wheat bread dough. *Journal of Cereal Science* **21**: 215–230.

Gellings, C. W. (2009). "Energy efficiency in fertilizer production and use". In: *Efficient User and Conservation of Energy - Volume II*, pp. 123–136.

Godfrey, D., Hawkesford, M. J., Powers, S. J., Millar, S., and Shewry, P. R. (2010). Effects of crop nutrition on wheat grain composition and end use quality. *Journal of Agricultural and Food Chemistry* **58**: 3012–3021.

Goesaert, H., Brijs, K., Veraverbeke, W. S., Courtin, C. M., Gebruers, K., and Delcour, J. A. (2005). Wheat flour constituents: How they impact bread quality, and how to impact their functionality. *Trends in Food Science and Technology* **16**: 12–30.

Gooding, M. J. (2009). "The wheat crop". In: *Wheat Chemistry and Technology.* Ed. by K Khan and P. Shewry. 4th Ed. St Paul: AACC, pp. 19–50.

Gooding, M. J., Ellis, R. H., Shewry, P. R., and Schofield, J. D. (2003). Effects of restricted water availability and increased temperature on the grain filling, drying and quality of winter wheat. *Journal of Cereal Science* **37**: 295–309.

Graybosch, R., Peterson, C. J., Moore, K. J., Stearns, M., and Grant, D. L. (1993). Comparative effects of wheat flour protein, lipid, and pentosan composition in relation to baking and milling quality. **70**: 95–101.

Gregersen, P. L., Holm, P. B., and Krupinska, K. (2008). Leaf senescence and nutrient remobilisation in barley and wheat. *Plant Biology* **10**: 37–49.

Hargin, K. D., Morrison, W. R., and Fulcher, R. G. (1980). Triglyceride deposits in the starchy endosperm of wheat. *Cereal Chemistry* **57**: 320–325.

He, J., Penson, S., Powers, S. J., Hawes, C., Shewry, P. R., and Tosi, P. (2013). Spatial patterns of gluten protein and polymer distribution in wheat grain. *Journal of Agricultural and Food Chemistry* **61**: 6207–6215.

Hennessy, K., Fawcett, R., Kirono, D., Jones, D., Bathols, J., Smith, M. S., Howden, M., Mitchell, C., and Plummer, N. (2008). An assessment of the impact of climate change on the nature and frequency of exceptional climatic events: drought exceptional circumstances, http://www.agriculture.gov.au/sitecollectiondocuments/agfood/drought/farmmanagementdeposits/csiro-bom-report-futu.

Henry, R. (1985). A Comparison of the non-starch carbohydrates in cereal grains. *Journal of the Science of Food and Agriculture* **36**: 1243–1253.

Holopainen, U. R. M., Wilhelmson, A., Home, S., Poutanen, K., and Shewry, P. R. (2012). Daylength effects on protein localisation affect water absorption in barley (Hordeum vulgare) grains. *Journal of the Science of Food and Agriculture* **92**: 2944–2951.

Hoseney, R. C. (1994). *Principles of cereal science and technology*. 2nd ed. Association of Cereal Chemists, Inc.

Hurkman, W. J. and Wood, D. F. (2011). High temperature during grain fill alters the morphology of protein and starch deposits in the starchy endosperm cells of developing wheat (*Triticum aestivum* L.) grain. *Journal of Agricultural and Food Chemistry* **59**: 4938–4946.

Hurkman, W. J., Tanaka, C. K., Vensel, W. H., Thilmony, R., and Altenbach, S. B. (2013). Comparative proteomic analysis of the effect of temperature and fertilizer on gliadin and glutenin accumulation in the developing endosperm and flour from *Triticum aestivum* I. cv. Butte 86. *Proteome Science* **11**: 8.

Hussain, I., Khan, M. A., and Khan, E. A. (2006). Bread wheat varieties as influenced by different nitrogen levels. *Journal of Zhejiang University SCIENCE B* **7**: 70–78.

IPCC (2001). "Working group I: The scientific basis". In: *IPCC Third Assessment Report: Climate Change 2001*. Ed. by J. Houghton, Y Ding, D. Griggs, M Noguer, P. van der Linden, X Dai, K Maskell, and C. Johnson. Cambridge University Press, pp. 185–237.

Islam, M. R., Haque, K. M. S., Akter, N., and Karim, M. A. (2014). Leaf chlorophyll dynamics in wheat based on SPAD meter reading and its relationship with grain yield. *Scientia Agriculturae* **8**: 13–18.

Jackson, E. A., Holt, L. M., and Payne, P. I. (1983). Characterisation of high molecular weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and the chromosomal localisation of their controlling genes. *Theoretical and Applied Genetics* **66**: 29–37.

Jenkins, G. J., Murphy, J. M., Sexton, D. M. H., Lowe, J. A., Jones, P., and Kilsby, C. G. (2009). *UK climate projections: Briefing report*, p. 59.

Jenner, C. F., Ugalde, T. D., and Aspinall, D. (1991). The physiology of starch and protein deposition in the endosperm of wheat. *Australian Journal of Plant Physiology* **18**: 211–226.

Kent, N. (1966). Subaleurone endosperm cells of high protein content. *Cereal Chemistry* **43**: 585–601.

Kent, N. L. and Evers, A. D. (1969). Variation in protein composition within endosperm of hard wheat. *Cereal Chemistry* **46**: 293–300.

Kettlewell, P. S., Sothern, R. B., and Koukkari, W. L. (1999). U.K. wheat quality and economic value are dependent on the North Atlantic Oscillation. *Journal of Cereal Science* **29**: 205–209.

Khatkar, B. S., Bell, A. E., and Schofield, J. D. (1995). The dynamic rheological properties of glutens and gluten sub-fractions from wheats of good and poor bread making quality. *Journal of Cereal Science* **22**: 29–44.

Kibbe, W. (2007). OligoCalc: An online oligonucleotide properties calculator. *Nucleic Acids Research* **35**: 43–46.

Kichey, T., Hirel, B., Heumez, E., Dubois, F., and Le Gouis, J. (2007). In winter wheat (*Triticum aestivum* L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. *Field Crops Research* **102**: 22–32.

Kindred, D. R., Verhoeven, T. M. O., Weightman, R. M., Swanston, J. S., Agu, R. C., Brosnan, J. M., and Sylvester-Bradley, R. (2008). Effects of variety and fertiliser nitrogen on alcohol yield, grain yield, starch and protein content, and protein composition of winter wheat. *Journal of Cereal Science* **48**: 46–57.

Koga, S., Böcker, U., Moldestad, A., Tosi, P., Shewry, P. R., Mosleth, E., and Uhlen, A. K. (2015). Influence of temperature on the composition and polymerization of gluten proteins during grain filling in spring wheat (*Triticum aestivum* L.) *Journal of Cereal Science* **65**: 1–8.

Kreis, M., Shewry, P. R., Forde, B. G., Forde, J., and Miflin, B. (1985). Structure and evolution of seed storage proteins and their genes with particular reference to those of wheat, barley and rye. *Oxford Surveys of Plant Molecular and Cell Biology* **2**: 253–317.

Lam, H. M., Coschigano, K. T., Oliveria, I. C., Melo-Oliveira, R., and Coruzzi, G. M. (1996). The molecular-genetics of nitrogen assimilation into amino acids in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 569–593.

Lásztity, R. (1996). *The chemistry of cereal proteins*. Ed. by R. Lásztity. 2nd Ed. Boca Raton: CRC Press.

Lobell, D. B. and Field, C. B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters* **2**: 014002.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., Van Ommen, T., Smith, A., and Elkins, J. (2006). Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP. *Geophysical Research Letters* **33**: 2000–2003.

MacRitchie, F. (1987). Evaluation of contributions from wheat protein fractions to dough mixing and breadmaking. *Journal of Cereal Science* **6**: 259–268.

Malik, A. H., Prieto-Linde, M. L., Kuktaite, R., Andersson, A., and Johansson, E. (2011). Individual and interactive effects of cultivar maturation time, nitrogen regime and temper-

ature level on accumulation of wheat grain proteins. *Journal of the Science of Food and Agriculture* **91**: 2192–2200.

Margiotta, B., Urbano, M., Colaprico, G., Johansson, E., Buonocore, F., D'Ovidio, R., and Lafiandra, D. (1996). Detection of y-type subunit at the *Glu-a1* locus in some Swedish bread wheat lines. *Journal of Cereal Science* **23**: 203–211.

Mauritsen, T. and Pincus, R. (2017). Committed warming inferred from observations. *Nature Climate Change* **7**: 652–655.

McIntosh, R. A., Yamazaki, Y., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., and Xia, X. C. (2013). "Catalogue of gene symbols for wheat". In: *12th International Wheat Genetics Symposium*.

MEA (2005). Ecosystems and human well-being: Synthesis. *Island Press, Washington, DC*.

Miflin, B. J., Burgess, S. R., and Shewry, P. R. (1981). The development of protein bodies in the storage tissues of seeds: Subcellular separations of homogenates of barley, maize and wheat endosperms and of pea cotyledons. *Journal of Experimental Botany* **32**: 199–219.

Miglietta, F., Tanasescu, M., and Marica, A. (1995). The expected effects of climate change on wheat development. *Global Change Biology* **1**: 407–415.

Millar, S. (2003). The development of near infrared (NIR) spectroscopy calibrations for the predition of wheat and flour quality. *HGCA Project Report No. 310*.

Mitchell, R. A. C., Mitchell, V. J., Driscoll, S. P., Franklin, J., and Lawlor, D. W. (1993). Effects of increased CO2 concentration and temperature on growth and yield of winter wheat at two levels of nitrogen application. *Plant, Cell and Environment* **16**: 521–529.

Moldestad, A., Hoel, B., Böcker, U., Koga, S., Mosleth, E. F., and Uhlen, A. K. (2014). Temperature variations during grain filling obtained in growth tunnel experiments and its influence on protein content, polymer build-up and gluten viscoelastic properties in wheat. *Journal of Cereal Science* **60**: 406–413.

Monaghan, J. M., Snape, J. W., Chojecki, A. J. S, and Kettlewell, P. S (2001). The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica* **122**: 309–317.

Monostori, I., Árendás, T., Hoffman, B., Galiba, G., Gierczik, K., Szira, F., and Vágújfalvi, A. (2016). Relationship between SPAD value and grain yield can be affected by cultivar, environment and soil nitrogen content in wheat. *Euphytica* **211**: 103–112.

Montzka, S. A., Dutton, G. S., Yu, P., Ray, E., Portmann, R. W., Daniel, J. S., Kuijpers, L., Hall, B. D., Mondeel, D., Siso, C., Nance, J. D., Rigby, M., Manning, A. J., Hu, L., Moore, F., Miller, B. R., and Elkins, J. W. (2018). An unexpected and persistent increase in global emissions of ozone-depleting CFC-11. *Nature* **557**: 413–417.

Moore, K. L., Tosi, P., Palmer, R., Hawkesford, M. J., Grovenor, C. R. M., and Shewry, P. R. (2016). The dynamics of protein body formation in developing wheat grain. *Plant Biotechnology Journal* **14**: 1876–1882.

Morris, C. and Rose, S. (1996). "Wheat". In: *Cereal Grain Quality*. Ed. by R. Henry and P. Kettlewell. 1st Ed. London: Chapman and Hall, pp. 3–54.

Muccilli, V., Cunsolo, V., Saletti, R., Foti, S., Margiotta, B., Scossa, F., Masci, S., and Lafiandra, D. (2010). Characterisation of a specific class of typical low molecular weight

glutenin subunits of durum wheat by a proteomic approach. *Journal of Cereal Science* **51**: 134–139.

NABIM (2014). National Association of British and Irish Flour Millers. *http://www.nabim.org.uk/wheat/wheat-varieties/*.

Nakano, H., Morita, S., and Kusuda, O. (2008). Effect of nitrogen application rate and timing on grain yield and protein content of the bread wheat cultivar 'Minaminokaori' in southwestern Japan. *Plant Production Science* **11**: 151–157.

Nakano, S. I., Fujimoto, M., Hara, H., and Sugimoto, N. (1999). Nucleic acid duplex stability: Influence of base composition on cation effects. *Nucleic Acids Research* **27**: 2957–2965.

Nasehzadeh, M. and Ellis, R. H. (2017). Wheat seed weight and quality differ temporally in sensitivity to warm or cool conditions during seed development and maturation. *Annals of Botany* **120**: 479–493.

Okumoto, S. and Pilot, G. (2011). Amino acid export in plants: A missing link in nitrogen cycling. *Molecular Plant* **4**: 453–463.

Otteson, B. N., Mergoum, M., and Ransom, J. K. (2007). Seeding rate and nitrogen management effects on spring wheat yield and yield components. *Agronomy Journal* **99**: 1615–1621.

Parry, M. A. J. and Hawkesford, M. J. (2010). Food security: Increasing yield and improving resource use efficiency. *Proceedings of the Nutrition Society* **69**: 592–600.

Parry, M. L., Rosenzweig, C., Iglesias, A., Livermore, M., and Fischer, G. (2004). Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global Environmental Change* **14**: 53–67.

Pask, A. J. D., Sylvester-Bradley, R., Jamieson, P. D., and Foulkes, M. J. (2012). Quantifying how winter wheat crops accumulate and use nitrogen reserves during growth. *Field Crops Research* **126**: 104–118.

Payne, P. I., Holt, L. M., Burgess, S. R., and Shewry, P. R. (1986). Characterisation by two-dimensional gel electrophoresis of the protein components of protein bodies, isolated from the developing endosperm of wheat (*Triticum aestivum*). *Journal of Cereal Science* **4**: 217–223.

Peña, R. J. (2002). "Wheat for bread and other foods". In: *Bread Wheat Improvement and Production*. Ed. by B. Curtis, S Rajaram, and H Gómez Macpherson. Rome: Food and Agriculture Organization of the United Nations, pp. 483–494.

Pepler, S., Gooding, M. J., and Ellis, R. H. (2006). Modelling simultaneously water content and dry matter dynamics of wheat grains. *Field Crops Research* **95**: 49–63.

Posner, E. S. (2009). "Wheat flour milling". In: *Wheat Chemistry and Technology*. Ed. by K Khan and P. Shewry. 4th Ed. St Paul: AACC, pp. 119–152.

Prasad, P. V. V., Pisipati, S. R., Ristic, Z., Bukovnik, U., and Fritz, A. K. (2008). Impact of nighttime temperature on physiology and growth of spring wheat. *Crop Science* **48**: 2372–2380.

Raftery, A. E., Zimmer, A., Frierson, D. M. W., Startz, R., and Liu, P. (2017). Less than 2 °c warming by 2100 unlikely. *Nature Climate Change* **7**: 637–641.

Rahman, M. A., Chikushi, J., Yoshida, S., and Karim, A. J.M. S. (2009). Growth and yield components of wheat genotypes exposed to high temperature stress under control environment. *Bangladesh Journal of Agricultural Research* **34**: 361–372.

Rasheed, A., Xia, X., Yan, Y., Appels, R., Mahmood, T., and He, Z. (2014). Wheat seed storage proteins: Advances in molecular genetics, diversity and breeding applications. *Journal of Cereal Science* **60**: 11–24.

Ray, D. K., Mueller, N., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **8**: e66428.

Rieu, I. and Powers, S. J. (2009). Real-time quantitative RT-PCR: design, calculations, and statistics. *The Plant Cell* **21**: 1031–1033.

Robertson, G. P. and Vitousek, P. M. (2009). Nitrogen in agriculture: balancing the cost of an essential resource. *Annual Review of Environment and Resources* **34**: 97–125.

Robertson, M. J., Brooking, I. R., and Ritchie, J. T. (1996). Temperature response of vernalization in wheat: Modelling the effect on the final number of mainstem leaves. *Annals of Botany* **78**: 371–381.

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. A. M., Schmid, E., Stehfest, E., Yang, H., and Jones, J. W. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences* **111**: 3268–3273.

Ruijter, J. M., Ramakers, C., Hoogaars, W. M., Karlen, Y., Bakker, O., hoff, M. J. Van den, and Moorman, A. F. (2009). Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research* **37**: e45.

Sapirstein, H. D. and Fu, B. X. (1998). Intercultivar variation in the quantity of monomeric proteins, soluble and insoluble glutenin, and residue protein in wheat flour and relationships to breadmaking quality. *Cereal Chemistry* **75**: 500–507.

Savill, G. P., Michalski, S., Powers, S. J., Wan, Y., Tosi, P., Buchner, P., and Hawkesford, M. J. (2018). Temperature and nitrogen supply interact to determine protein distribution gradients in the wheat grain endosperm. *Journal of Experimental Botany* **69**: 3117–3126.

Sawyer, J. S. (1972). Man-made carbon dioxide and the "greenhouse" effect. *Nature* **239**: 23–26.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J., White, D, Hartenstein, V, Liceiri, K, Tomancak, P, and Cardona, A (2012). Fiji: An open source platform for biological image analysis. *Nature Methods* **9**: 676–682.

Schröder, J. J., Aarts, H. F. M., Berge, H. F. M. ten, Keulen, H. van, and Neeteson, J. J. (2003). An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. *European Journal of Agronomy* **20**: 33–44.

Shah, N. H. and Paulsen, G. M. (2003). Interaction of drought and high temperature on photosynthesis and grain-filling of wheat. *Plant and Soil* **257**: 219–226.

Shewry, P. R. (2009). Wheat. *Journal of Experimental Botany* **60**: 1537–1553.

Shewry, P. R., Tatham, A. S., Forde, J., Kreis, M., and Miflin, B. J. The classification and nomenclature of wheat gluten proteins: A reassessment. *Journal of Cereal Science* **4**: (), 97–106.

Shewry, P. R. and Tatham, A. S. (1997). Disulphide bonds in qheat gluten proteins. *Journal of Cereal Science* **25**: 207–227.

Shewry, P. R., Evers, A. D., Bechtel, D. B., and Abecassis, J. (2009a). "Development, structure and mechanical properties of the wheat grain". In: *Wheat Chemistry and Technology*. Ed. by K Khan and P. Shewry. 4th Ed. St Paul: AACC, pp. 51–96.

Shewry, P. R., D'Ovidio, R., Lafiandra, D., Jenkins, J., Mills, E., and Békés, F. (2009b). "Wheat grain proteins". In: *Wheat Chemistry and Technology*. Ed. by K Khan and P. Shewry. 4th Ed. St Paul: AACC, pp. 223–298.

Simmonds, N. (1995). The relation between yield and protein in cereal grain. *Journal of the Science of Food and Agriculture* **67**: 309–315.

Singh, J., Blundell, M., Tanner, G., and Skerritt, J. H. (2001a). Albumin and globulin proteins of wheat flour: Immunological and N-terminal sequence characterisation. *Journal of Cereal Science* **34**: 85–103.

Singh, J. and Skerritt, J. H. (2001b). Chromosomal control of albumins and globulins in wheat grain assessed using different fractionation procedures. *Journal of Cereal Science* **33**: 163–181.

Singh, N. K. (1987). Solubility behavior, synthesis, degradation and subcellular location of a new class of disulphide-linked proteins in wheat endosperm. *Australian Journal of Plant Physiology* **14**: 245–252.

Slafer, G. A., Savin, R., and Sadras, V. O. (2014). Coarse and fine regulation of wheat yield components in response to genotype and environment. *Field Crops Research* **157**: 71–83.

Snyder, C. S., Bruulsema, T. W., Jensen, T. L., and Fixen, P. E. (2009). Review of greenhouse gas emissions from crop production systems and fertilizer management effects. *Agriculture, Ecosystems and Environment* **133**: 247–266.

Sofield, I., Evans, L. T., Cook, M. G., and Wardlaw, I. F. (1977). Factors influencing the rate and duration of grain filling in wheat. *Australian Journal of Plant Physiology* **4**: 785–797.

Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M, and Miller, H. L. (2007). Climate change 2007: The physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC)*, 996.

Sosulski, F. W. and Imafidon, G. I. (1990). Amino acid composition and nitrogen-to-protein conversion factors for animal and plant foods. *Journal of Agricultural and Food Chemistry* **38**: 1351–1356.

Southgate, D. (2009). Population growth, increases in agricultural production and trends in food prices. *The Electronic Journal of Sustainable Development* **1**: (3).

Spiertz, J. H. J., Hamer, R. J., Xu, H., Primo-Martin, C., Don, C., and Putten, P. E. L. van der (2006). Heat stress in wheat (*Triticum aestivum* L.): Effects on grain growth and quality traits. *European Journal of Agronomy* **25**: 89–95.

Sreeramulu, G. and Singh, N. K. (1997). Genetic and biochemical characterization of novel low molecular weight glutenin subunits in wheat (*Triticum aestivum* L.) *Genome* **40**: 41–48.

Stone, B. and Morell, M. K. (2009). "Carbohydrates". In: *Wheat Chemistry and Technology*. Ed. by K Khan and P. R. Shewry. 4th Ed. St Paul: AACC, pp. 299–362.

Tao, Z., Chang, X., Wang, D., Wang, Y., Ma, S., Yang, Y., and Zhao, G. (2018). Effects of sulfur fertilization and short-term high temperature on wheat grain production and wheat flour proteins. *Crop Journal* **6**: 413–425.

Tashiro, T. and Wardlaw, I. (1990). The response to high temperature shock and humidity changes prior to and during the early stages of grain development in wheat. *Australian Journal of Plant Physiology* **17**: 551.

Tatham, A. S., Gilbert, S. M., Fido, R. J., and Shewry, P. R. (2000). Extraction, separation, and purification of wheat gluten proteins and related proteins of barley, rye, and oats. *Methods in Molecular Medicine* **41**: 55–73.

Taub, D. R. and Wang, X. (2008). Why are nitrogen concentrations in plant tissues lower under elevated CO2? A critical examination of the hypotheses. *Journal of Integrative Plant Biology* **50**: 1365–1374.

Thorne, G. N. and Wood, D. W. (1987). Effects of radiation and temperature on tiller survival, grain number and grain yield in winter wheat. *Annals of Botany* **59**: 413–426.

Tilman, D. (1999). Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences* **96**: 5995–6000.

Tilman, D., Fargione, J., Wolff, B., Antonio, C. D., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W. H., Simberloff, D., and Swackhamer, D. (2001). Forecasting agriculturally driven environmental change. *Science* **292**: 281–284.

Tosi, P., Gritsch, C. S., He, J., and Shewry, P. R. (2011). Distribution of gluten proteins in bread wheat (*Triticum aestivum*) grain. *Annals of Botany* **108**: 23–35.

Tottman, D. R. (1987). The decimal code for the growth stages of cereals, with illustrations. *Annals of Applied Biology* **110**: 441–454.

Ugalde, T. D. and Jenner, C. F. (1990a). Substrate gradients and regional patterns of dry matter deposition within developing wheat endosperm. I. Carbohydrates. *Functional Plant Biology* **17**: 377–394.

Ugalde, T. D. and Jenner, C. (1990b). Substrate gradients and regional patterns of dry matter deposition within developing wheat endosperm. II. Amino acids and protein. *Australian Journal of Plant Physiology* **17**: 395–406.

United Nations (2017). World population prospects: The 2017 revision, key findings and advance tables. *Working Paper No. ESA/P/WP/248*.

Uthayakumaran, S., Newberry, M., Keentok, M., Stoddard, F. L., and Békés, F. (2000). Basic rheology of bread dough with modified protein content and glutenin-to-gliadin ratios. *Cereal Chemistry* **77**: 744–749.

Veraverbeke, W. S. and Delcour, J. A. (2002). Wheat protein composition and properties of wheat glutenin in relation to breadmaking functionality. *Critical Reviews in Food Science and Nutrition* **42**: 179–208.

Victor, D. G., Akimoto, K., Kaya, Y., Yamaguchi, M., Cullenward, D., and Hepburn, C. (2017). Prove Paris was more than paper promises. *Nature* **548**: 25–27.

Wan, Y., Shewry, P. R., and Hawkesford, M. J. (2013). A novel family of γ-gliadin genes are highly regulated by nitrogen supply in developing wheat grain. *Journal of Experimental Botany* **64**: 161–168.

Wan, Y., Gritsch, C. S., Hawkesford, M. J., and Shewry, P. R. (2014). Effects of nitrogen nutrition on the synthesis and deposition of the ω-gliadins of wheat. *Annals of Botany* **113**: 607–615.

Wang, Y. G., Khan, K., Hareland, G., and Nygard, G. (2007). Distribution of protein composition in bread wheat flour mill streams and relationship to breadmaking quality. *Cereal Chemistry* **84**: 271–275.

Wang, Y. P. and Gifford, R. M. (1995). A model of wheat grain growth and its applications to different temperature and carbon dioxide levels. *Australian Journal of Plant Physiology* **22**: 843–855.

Wheeler, T. and Braun, J. von (2013). Climate change impacts on global food security. *Science* **341**: 508–513.

Yahata, E., Maruyama-Funatsuki, W., Nishio, Z., Yamamoto, Y., Hanaoka, A., Sugiyama, H., Tanida, M., and Saruyama, H. (2006). Relationship between the dough quality and content of specific glutenin proteins in wheat mill streams, and its application to making flour suitable for instant Chinese noodles. *Bioscience, Biotechnology, and Biochemistry* **70**: 788–797.

Yang, F., Jørgensen, A. D., Li, H., Søndergaard, I., Finnie, C., Svensson, B., Jiang, D., Wollenweber, B., and Jacobsen, S. (2011). Implications of high-temperature events and water deficits on protein profiles in wheat (*Triticum aestivum* L. cv. Vinjett) grain. *Proteomics* **11**: 1684–1695.

Yasunaga, T., Bushuk, W., and Irvine, G. (1968). Gelatinization of starch during bread-baking. *Cereal Chemistry* **45**: 269–279.

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stage of cereals. *Weed Research* **14**: 415–421.

Zhang, H. B., Dai, H. C., Lai, H. X., and Wang, W. T. (2017). U.S. withdrawal from the Paris Agreement: Reasons, impacts, and China's response. *Advances in Climate Change Research* **8**: 220–225.

Zhang, P., He, Z., Chen, D., Zhang, Y., Larroque, O., and Xia, X. (2007a). Contribution of common wheat protein fractions to dough properties and quality of northern-style Chinese steamed bread. *Journal of Cereal Science* **46**: 1–10.

Zhang, P., He, Z., Zhang, Y., Xia, X., and Liu, J. (2007b). Pan bread and chinese white salted noodle qualities of Chinese winter wheat cultivars and their relationship with gluten protein fractions. *Cereal Chemistry* **84**: 370–378.

Zhou, Q., Li, X., Yang, J., Zhou, L., Cai, J., Wang, X., Dai, T., Cao, W., and Jiang, D (2018). Spatial distribution patterns of protein and starch in wheat grain affect baking quality of bread and biscuit. *Journal of Cereal Science* **79**: 362–369.

Appendix A: Experimental protocols

A.1. Microscopy sample fixation, dehydration, and embedding

Materials

Chemicals:

- 0.1M Sorenson's phosphate buffer (pH 7.4)
- 4% paraformaldehyde + 2.5% gluteraldehyde fixative in 0.1M phosphate buffer
- Ethanol series from 10% to 100% dry ethanol in increments of 10%
- LR white resin, medium grade

Equipment:

- Double edge razor blades
- Glass vials
- Pencil and card for sample labels
- Polypropelene embedding capsules
- 55°C nitrogen-gas-filled oven
- Vacuum chamber and pump

Protocol

All work with fixative or LR white resin must be conducted in a fume cabinet.

1. Prepare 0.1M Sorensen's phosphate buffer (pH 7.4).
   (a) Mix 9.5ml of 0.2M NaH_2PO_4 with 40.5ml of 0.2M Na_2HPO_4.
   (b) Bring up to 100ml with RO water.
   (c) Test to ensure pH is at 7.4.
2. Prepare 4% paraformaldehyde + 2.5% gluteraldehyde fixative in fume cabinet.
   (a) Dissolve 4g of paraformaldehyde in 50ml of 0.1M Sorensen's phosphate buffer (pH 7.4) in a water bath at 70°C. Invert occasionally until solution goes completely clear. Cool on ice.
   (b) Add 10ml of 25% gluteraldehyde solution.

(c) Bring up to 100ml with 0.1M Sorensen's phosphate buffer (pH 7.4).

3. Ensure fixative and buffer are at room temperature (RT) before using.

4. Place a drop of fixative on the grain to be sampled on a piece of clean dental wax or ceramic tile, and cut 1mm sections from the wheat grain with a sterilised double edge razor blade using a sliding motion to avoid crushing the grain.

5. Place tissue sections immediately in a glass vial containing fixative. Use a mild vacuum to expel air from the sample and to improve infiltration of the fixative into the grain tissue. Apply vacuum for about 2 minutes before releasing. Repeat three times. Place vials on a rotator at RT for 3–5 hours.

6. Wash specimens three times in 0.1M Sorensen's phosphate buffer (pH 7.4) for 30 minutes in each wash. Ensure samples are not exposed to air to prevent air infiltrating the sample. Samples can be stored at 4°C prior to futher processing.

7. Dehydrate samples in a graded ethanol series with samples at RT on a rotator:
   (a) 10% ethanol at RT for 1 hour.
   (b) 20% ethanol at RT for 1 hour.
   (c) 30% ethanol at RT for 1 hour.
   (d) 40% ethanol at RT for 1 hour.
   (e) 50% ethanol at RT for 1 hour. Repeat.
   (f) 60% ethanol at RT for 1 hour. Repeat.
   (g) 70% ethanol at 4°C overnight.
   (h) 80% ethanol at RT for 1 hour. Repeat.
   (i) 90% ethanol at RT for 1 hour. Repeat.
   (j) 100% dry ethanol at RT for 1 hour. Repeat twice.

8. Infiltrate samples with resin in increasing concentrations of medium grade LR white resin. As with the fixative, use a mild vacuum to aid infiltration of the resin. Ensure resin is brought to RT prior to opening. Samples can be stored at 4°C overnight between steps.
   (a) Ethanol : LR white at 4 : 1 for at least 6 hours.
   (b) Ethanol : LR white at 3 : 2 for at least 6 hours.
   (c) Ethanol : LR white at 2 : 3 for at least 6 hours.
   (d) Ethanol : LR white at 1 : 4 for at least 6 hours.
   (e) Pure LR white resin for 1 hour. Repeat twice.

(f) Two changes of pure LR white resin per day for at least 5 days. Place samples in rotator at RT during the day, and store at 4°C overnight.

9. Fill labeled embedding capsules with fresh resin. Place and orientate samples within the capsules. Polymerise samples in the oven at 55°C for 16–24 hours in an oxygen-free environment. Allow samples to cool prior to sectioning.

A.2. Protein extraction for SDS-PAGE

Materials

Chemicals:

- 50% propan-1-ol + 2% DTT
- Total loading buffer:
  - 50mM Tris-HCL (pH 6.8)
  - 2% (w/v) SDS
  - 10% (v/v) glycerol
  - 0.1% (w/v) Bromophenol blue
  - 200mM DTT
- Liquid N_2 (for milling)

Equipment:

- SPEC SamplePrep 6870 Freezer/Mill®
- Edwards Modulyo® freeze drier

Protocol

1. Suspend 10mg of flour in 150µl of 50% propan-1-ol + 2% DTT by vortexing.
2. Shake at 50°C for 45 minutes.
3. Centrifuge at maximum speed for 15 minutes.
4. Remove supernatent to new 1.5ml micro-centrifuge tube.
5. Resuspend pellet in 150µl of 50% propan-1-ol + 2% DTT.
6. Shake at 50°C for 45 minutes.
7. Centrifuge at maximum speed for 15 minutes.
8. Remove and combine with previous supernatent from step 4.
9. Freeze-dry supernatent.
10. Add 150µl of total loading buffer.
11. Heat sample to 90°C for 3 minutes.
12. Centrifuge at maximum speed for 15 minutes.

13. Supernatent contains wheat gluten protein extract. Store at -20°C, and repeat steps 11 and 12 prior to running frozen samples on a gel.

A.3. **Running SDS-PAGE gel**

Materials

Chemicals:

- MES running buffer
- Coomassie gel stain:
  - 0.1% (w/v) Coomassie Brilliant Blue (R-250)
  - 40% (v/v) methanol
  - 10% (w/v) TCA
  - Made up to 100% with distilled water (dH_2O)
- De-stain solution:
  - 10% TCA
  - Made up to 100% with dH_2O

Equipment:

- Bolt® Mini Gel Tank
- Pre-cast 17-well Bolt® 8% gradient Bis-Tris gels

Protocol

1. Rinse pre-case gels with distilled water, and wash wells with MES running buffer.
2. Load 1–10µl of protein sample (depending on concentration) per lane.
3. Run gel for 30 minutes at a constant current of 20mA to settle sample in bottom of well.
4. Run gel for a further 330 minutes at a constant current of 50mA.
5. Remove gel from cassette and stain in Coomassie stain overnight on a rocking shaker.
6. Destain gel with 10% TCA on a rocking shaker with a small strip of foam or paper towel in the container to absorb the stain.
7. Remove gel from the stain when the background is clear, but protein bands remain brightly stained.
8. Image the gel.

A.4. Protein extraction for SE-HPLC

Materials

Chemicals:

- Protein extraction buffer:
  - 2% (w/v) SDS
  - 0.1M NaH_2PO_4 phosphate buffer
  - pH 6.9 with HCl

Equipment:

- Ultrasonic disintegrator fitted with 3mm exponential tip
- 2ml disposable syringes
- 0.45µm syringe filters (Gilson, UK)
- 2ml glass vials with silicone/PTFE resealable caps (Thermo Scientific, UK)

Protocol

1. Measure out 16.5mg of freeze-dried whole-grain flour into a 2ml micro-centrifuge tube.
2. Prepare SDS protein extraction buffer and set pH to 6.9 with HCl.
3. Add 1.5ml of SDS protein extraction buffer to each flour sample.
4. Sonicate each sample for 45 seconds, inserting the microtip into the sample.
5. Centrifuge at maximum speed for 30 minutes.
6. Aliquot supernatent containing extracted protein into 2ml glass vials using disposable syringe with 0.45µm filter.

A.5. **RNA extraction**

Materials

Chemicals:

- Molecular grade phenol
- Chloroform:IAA (24:1)
- Phenol:chloroform:IAA (25:24:1)
- RNA extraction buffer:
  - 0.1M Tris-HCl
  - 0.1M LiC
  - 1% (w/v) SDS
  - 10mM EDTA
  - pH 8.0
- DNAse treatment (per sample):
  - 15µl 10x DNAse buffer
  - 127µl H_2O (DEPC treated)
  - 8µl DNAse (RNAse-free)
- 3M Na-acetate, pH 5.2 (acetic acid)
- 4M LiCl
- 70% EtOH
- Promega DNAse (RNAse-free), and 10x buffer
- 100% EtOH
- DEPC treated H_2O

With the exception of the RNA extraction buffer, the H_2O used to make all of the above are treated with DEPC prior to autoclaving to destroy any RNAse activity—0.1% DEPC v/v, incubated at RT for several hours, and autoclaved prior to use.

Equipment:

- SPEC SamplePrep 6870 Freezer/Mill®
- Refrigerated centrifuge (4°C)

- 80°C water bath
- Thermo Scientific NanoDrop 2000c spectrophotometer

Protocol

All work with phenol or chloroform must be conducted in a fume cabinet.

Day 1:

1. Prepare phenol:extraction buffer at a ratio of 8:12, and heat to 80°C in water bath.
2. Transfer approximately 0.5ml of frozen ground sample into a pre-cooled 2ml micro-centrifuge tube.
3. Add 1ml of hot (80°C) phenol:extraction buffer to each sample and vortex until material is completely thawed and homogenised. Vortex for a further 30 seconds.
4. Add 0.5ml of chloroform:IAA and vortex for 30 seconds.
5. Centrifuge at maximum speed for 5 minutes at 4°C.
6. Transfer aqueous phase to fresh 2ml micro-centrifuge tube. If interphase is large, repeat steps 3–5.
7. Add 1ml of chloroform/IAA and vortex for 30 seconds.
8. Centrifuge at maximum speed for 5 minutes at 4°C.
9. Transfer aqueous phase to fresh 1.5ml micro-centrifuge tube.
10. Measure volume, and add an equal volume +20µl of 4M LiCl.
11. Mix well and incubate at 4°C overnight.

Day 2:

1. Centrifuge at maximum speed for 20 minutes at 4°C.
2. Discard the supernatent and wash the pellet with 1ml of 70% ethanol.
3. Centrifuge at maximum speed for 5 minutes at 4°C.
4. Discard the supernatent and allow pellet to dry.
5. Add 150µl of DNAse treatment, and dissolve pellet on ice for 30–40 minutes.
6. Shake at 37°C for 30 minutes.
7. Add 150µl of H_2O.

8. Add 300µl of chloroform:IAA and vortex for 30 seconds.
9. Centrifuge at maximum speed for 5 minutes at 4°C.
10. Transfer aqueous phase to fresh 1.5ml micro-centrifuge tube.
11. Add 300µl of phenol:chloroform:IAA and vortex for 30 seconds.
12. Centrifuge at maximum speed for 5 minutes at 4°C.
13. Transfer aqueous phase to 1.5ml micro-centrifuge from step 10.
14. Add 1/10 volume of 3M NaOAc (pH 5.2) and 2.5x volume of ethanol.
15. Mix well and incubate at -20°C overnight.

Day 3:

1. Centrifuge at maximum speed for 20 minutes at 4°C.
2. Discard the supernatent and wash with 1ml of 70% ethanol.
3. Centrifuge at maximum speed for 5 minutes at 4°C.
4. Discard the supernatent and allow pellet to dry.
5. Dissolve pellet on ice in H_2O—30–150µl, depending on size of pellet.
6. Shake at 37°C for 5 minutes.
7. Centrifuge at maximum speed for 5 minutes at 4°C.
8. Transfer supernatent into a fresh 1.5ml micro-centrifuge tube.
9. Measure the concentration of RNA at 260nm using NanoDrop 2000c spectrophotometer.
10. Check quality of 1µg of RNA by TAE-agarose (1%) electrophoresis.

A.6. cDNA synthesis

Materials

Chemicals:

- 2µg of extracted RNA
- 10µM dT-AP primer
- 5X first strand buffer (Invitrogen^{TM})
- 0.1M DTT
- 10mM dNTP mix
- Superscript^{TM} III reverse transcriptase (Invitrogen^{TM})
- H_2O treated with DEPC

Equipment:

- Bio-Rad MJ Research deoxyribonucleic acid (DNA) Engine thermal cycler

Protocol

1. Prepare and label 0.2ml micro-centrifuge tubes.
2. Add enough DEPC-treated H_2O to make a final volume to 13µl after primers and RNA have been added.
3. Add 1µl of dT-AP primer.
4. Add 2µg of extracted RNA.
5. Incubate for 7 minutes at 70°C in PCR machine. Chill immediately on ice.
6. Centrifuge at max speed to collect contents of the micro-centrifuge tubes.
7. Add the following as a mix:

   4µl of 5x first strand buffer

   1µl of 0.1M DTT

   1µl 10mM dNTP mix

   1µl Superscript^{TM} III reverse transciptase
8. Mix gently and centrifuge to collect contents of the micro-centrifuge tubes.
9. Incubate in PCR machine as follows:

5 minutes at 22°C

2 hours at 50°C

15 minutes at 70°C

Hold at 10 °C

10. Centrifuge at max speed to collect contents of the micro-centrifuge tubes. Store cDNA at -20°C.

A.7. Quantitive PCR (qPCR)

Materials

Chemicals:

- 1.1µl sample cDNA
- 10µM sense primer
- 10µM anti-sense primer
- 100x ROX internal reference dye (Sigma-Aldrich)
- SYBR® Green JumpStart™ Taq ReadyMix™ (Sigma-Aldrich)
- H_2O treated with DEPC

Equipment:

- 96-well PCR plates with qPCR transparent sealing film
- Applied Biosystems™ 7500 Real Time PCR System

Protocol

1. Prepare and label 1.5ml micro-centrifuge tubes.
2. Create a master mix containing the following per sample:

   0.7µl of sense primer (10µM)

   0.7µl of anti-sense primer (10µM)

   0.03µl of 100x ROX internal reference dye

   11.22µl of H_2O

   13.75µl SYBR® Green JumpStart™ Taq ReadyMix™
3. Add 26.4µl of the master mix to each of the prepared 1.5ml tubes.
4. Add 1.1µl of cDNA to each tube. Mix well, and collect by centrifugation.
5. Transfer 25.1µl into each well of a white 96-well PCR plate. Seal plate with transparent seal, and centrifuge to collect contents.
6. Run plate in the 7500 real-time PCR machine with the following program:

   2 minutes at 50°C

   10 minutes at 95°C

Repeat the following two steps 41 times:

15 seconds at 95°C

1 minute at 60°C

15 seconds at 95°C

15 seconds at 60°C

7. Download results.

Appendix B: Microscopy imaging coordinate calculator app

Since the focal length of the light-microscope is very small at 20x magnification, any variation in the distance of the sample from the microscope lens results in the image going out of focus. This effect is significant when imaging large sections such as wheat grain. To allow for the capture of in-focus images, four overlapping images of each grain must be taken, with the microscopy re-focussed before each image is captured. These four images must then combined into a single image prior to analysis.

The following Python 3 GUI application calculates the coordinates required to take four overlapping images. It takes coordinates for the top left and bottom right of the area to be imaged as inputs, and outputs the coordinates of the top left and bottom right coordinates required for the four overlapping images to be captured. Must be compiled with py2exe prior to use.

```
From Tkinter import *

# Function to calculate output coordinates from input coordinates
def calculate(*args):
    try:
      TLxDiff.set(TLx.get()+(((BRx.get()-TLx.get())/2.0)*1.25))
      TLyDiff.set(TLy.get()+(((BRy.get()-TLy.get())/2.0)*1.25))
      BRxDiff.set(BRx.get()-(((BRx.get()-TLx.get())/2.0)*1.25))
      BRyDiff.set(BRy.get()-(((BRy.get()-TLy.get())/2.0)*1.25))
    except ValueError:
        pass

# Open GUI window
root = Tk()
root.title("Co-ordinates Calculator")
root.geometry("360x500")

# Assign input and output variables
TLx = IntVar()
TLy = IntVar()
BRx = IntVar()
BRy = IntVar()

TLxDiff = IntVar()
TLyDiff = IntVar()
```

```
BRxDiff = IntVar()
BRyDiff = IntVar()

# Draw input text boxes
TLx_entry0 = Entry(root, width=12, textvariable=TLx)
TLx_entry0.grid(column=2, row=1)
TLy_entry0 = Entry(root, width=12, textvariable=TLy)
TLy_entry0.grid(column=3, row=1)
BRx_entry0 = Entry(root, width=12, textvariable=BRx)
BRx_entry0.grid(column=2, row=2)
BRy_entry0 = Entry(root, width=12, textvariable=BRy)
BRy_entry0.grid(column=3, row=2)

# Draw of labels for four input boxes.
Label(root, text="x").grid(column=2, row=0)
Label(root, text="y").grid(column=3, row=0)
Label(root, text="Top Left").grid(column=1, row=1)
Label(root, text="Bottom Right").grid(column=1, row=2)
Label(root, text=" ").grid(column=1, row=3)

# Display image 1 (top left) output
Label(root, text="Image 1 (Top Left)").grid(column=1, row=4)
TLx_entry1 = Entry(root, width=12, textvariable=TLx)
TLx_entry1.grid(column=2, row=5)
TLy_entry1 = Entry(root, width=12, textvariable=TLy)
TLy_entry1.grid(column=3, row=5)
BRx_entry1 = Entry(root, width=12, textvariable=TLxDiff)
BRx_entry1.grid(column=2, row=6)
BRy_entry1 = Entry(root, width=12, textvariable=TLyDiff)
BRy_entry1.grid(column=3, row=6)

# Display image 2 (top right) output
Label(root, text="Image 2 (Top Right)").grid(column=1, row=7)
TLx_entry2 = Entry(root, width=12, textvariable=BRxDiff)
TLx_entry2.grid(column=2, row=8)
TLy_entry2 = Entry(root, width=12, textvariable=TLy)
TLy_entry2.grid(column=3, row=8)
BRx_entry2 = Entry(root, width=12, textvariable=BRx)
BRx_entry2.grid(column=2, row=9)
BRy_entry2 = Entry(root, width=12, textvariable=TLyDiff)
```

```
BRy_entry2.grid(column=3, row=9)

# Display image 3 (bottom left) output
Label(root, text="Image 3 (Bottom Left)").grid(column=1, row=10)
TLx_entry3 = Entry(root, width=12, textvariable=TLx)
TLx_entry3.grid(column=2, row=11)
TLy_entry3 = Entry(root, width=12, textvariable=BRyDiff)
TLy_entry3.grid(column=3, row=11)
BRx_entry3 = Entry(root, width=12, textvariable=TLxDiff)
BRx_entry3.grid(column=2, row=12)
BRy_entry3 = Entry(root, width=12, textvariable=BRy)
BRy_entry3.grid(column=3, row=12)

# Display image 4 (bottom right) output
Label(root, text="Image 4 (Bottom Right)").grid(column=1, row=13)
TLx_entry4 = Entry(root, width=12, textvariable=BRxDiff)
TLx_entry4.grid(column=2, row=14)
TLy_entry4 = Entry(root, width=12, textvariable=BRyDiff)
TLy_entry4.grid(column=3, row=14)
BRx_entry4 = Entry(root, width=12, textvariable=BRx)
BRx_entry4.grid(column=2, row=15)
BRy_entry4 = Entry(root, width=12, textvariable=BRy)
BRy_entry4.grid(column=3, row=15)

# Create 'padding' around the edge of the window
for child in root.winfo_children(): child.grid_configure(padx=11, pady=5)

# Set the initial focus of the window to the Top Left X coordinate input box
TLx_entry0.focus()

# Call 'calculate' function whenever a key is pressed
root.bind('<Key>', calculate)

root.mainloop()
```

Appendix C: Python toolbox for ArcGIS

Python toolbox for use in ArcCatalog for the protein concentration gradient and protein body size-distribution analysis of wheat grain images. As primary inputs takes a TIFF image file of the stained wheat grain, a shapefile (.shp) of the outline of the wheat grain drawn in ArcMap, an image classification signature file (.gsg) generated with training samples defined in ArcMap, and the treatment name/code as a text input. Inputs are also taken for the number of samples within the signature file that represent area of interest (default = 10), the number of zones to be drawn (default = 5), an optional input file for the widths of the zones to be drawn (.txt), and a scalebar.txt file that can be used to override the default scaling factor for converting from arbitrary units to micrometers (two line file: first line is length of scalebar in um, second line is length of scalebar in pixels).

Toolbox produces two outputs as csv files: the result of the protein concentration gradient analysis (treatment_zones.csv), and of the protein body size-distribution analysis (treatment_spatial.csv). Additionally, an output of the maximum grain width is produced for each analysis, but not permenantly stored, and is overwritten by subsequent analyses.

A second "RescalingBatch" toolbox is required to rescale input images to a 1×1 cell size. This is required to prevent inaccuracies in the conversion of measurements from pixels to micrometers.

```
# ArcPy toolbox for use with ArcGIS. Written by Adam Michalski and modified
# by George Savill

import os, sys, string, arcpy
import numpy as np
import glob
from arcpy import env
from arcpy.sa import *
arcpy.env.pyramid = "NONE" # Improves processing time
arcpy.env.overwriteOutput=True
arcpy.env.cellSize = 1 # Set the output raster cell size

class Toolbox(object):
    def __init__(self):
        """Define the toolbox (the name of the toolbox is the name of the
            .pyt file)."""
        self.label = "Extracting Proteins"
        self.alias = ""
```

```
        # List of tool classes associated with this toolbox
        self.tools = [Protein,RescalingBatch]

# Main class for analysis of microscopy images
# Calculates protein concentration gradient and protein body
    size-distribution
# data
class Protein(object):
    def __init__(self):
        self.label = 'Extract proteins from the grain'
        self.canRunInBackground = False

    def getParameterInfo(self):
        '''Set up the parameters and return the list of parameter objects.'''

        # Input grain raster image
        param0 = arcpy.Parameter()
        param0.name = 'tifFile'
        param0.displayName = '1. Input grain TIF file:'
        param0.parameterType = 'Required'
        param0.direction = 'Input'
        param0.datatype = 'DEFile'
        param0.filter.list = ['tif', 'tiff', 'TIF']

        # Input image classification signature file
        param1 = arcpy.Parameter()
        param1.name = 'signature'
        param1.displayName = '2. Input signature file:'
        param1.parameterType = 'Required'
        param1.direction = 'Input'
        param1.datatype = 'DEFile'
        param1.filter.list = ['gsg']

      # Input number of samples representing protein (default = 10)
        param2 = arcpy.Parameter()
        param2.name = 'number_of_samples'
        param2.displayName = '3. Input number of samples:'
        param2.parameterType = 'Required'
        param2.direction = 'Input'
        param2.datatype = 'GPLong'
```

```
        param2.value = 10 # Default value

    # Input number of zones to be drawn (default = 5)
        param3 = arcpy.Parameter()
        param3.name = 'number_of_zones'
        param3.displayName = '4. Input number of zones:'
        param3.parameterType = 'Required'
        param3.direction = 'Input'
        param3.datatype = 'GPLong'
        param3.value = 5 # Default value

    # Input text file with zone distances (optional)
        param4 = arcpy.Parameter()
        param4.name = 'textZones'
        param4.displayName = '5. Input distance zones text file:'
        param4.parameterType = 'Optional'
        param4.direction = 'Input'
        param4.datatype = 'DETextfile'
        param4.filter.list = ['txt']

        # Input outline of grain shapefile
        param5 = arcpy.Parameter()
        param5.name = 'Borders'
        param5.displayName = '6. Select shapefile (*.shp) with grain
            borders:'
        param5.parameterType = 'Required'
        param5.direction = 'Input'
        param5.datatype = 'Shapefile'
        param5.filter.list = ['shp']

    # Input text description field
        param6 = arcpy.Parameter()
        param6.name = 'textDescription'
        param6.displayName = '7. Text description column:'
        param6.parameterType = 'Optional'
        param6.direction = 'Input'
        param6.datatype = 'String'

        return [param0,param1,param2,param3,param4,param5,param6]
```

```
def isLicensed(self):
    """Prevent the tool from running if the Spatial Analyst extension is
       not available."""
    if arcpy.CheckExtension('Spatial') == 'Available':
        return True # The tool can be executed.
    else:
        return False # The tool can not be executed.

def updateParameters(self, parameters):
    return

def updateMessages(self, parameters):
    return

def execute(self, parameters, messages):
  # Assigning input parameters to variable names
  grainTIF = parameters[0].valueAsText
  signatureFILE = parameters[1].valueAsText
  numberSAMPLES = parameters[2].value
  numbZONES = parameters[3].value
  txtFILEdistances = parameters[4].valueAsText
  shpBORDER = parameters[5].valueAsText
  description = parameters[6].valueAsText

  arcpy.env.overwriteOutput=True

  # Calculate Maximum Likelihood (ML) of pixels in input raster image
     using
  # image classification signature file
  mlRASTER = MLClassify(grainTIF, signatureFILE)

  # Con ML
  # Identifying which pixels were identified as protein by ML
  # classification (according to numberSAMPLES)
  numberSAMPLESstring = str(numberSAMPLES)
  where_clause = "VALUE >= 1 AND VALUE <= " + numberSAMPLESstring
  protCON = Con(mlRASTER, 1, "", where_clause)

  # RasterToPolygon
  folderIN = os.path.dirname(grainTIF)
```

```
polySHP = folderIN + "/" + "shape.shp"
arcpy.RasterToPolygon_conversion(protCON, polySHP, "NO_SIMPLIFY")

# Constants for conversion to micrometers (um).
constantLINEAR = 0.320595024365221851756860733521 41 # for distance
constantAREA = 0.10278116964773719279201741905662 # for area

# Calculating scaling for distance (constantLINEAR) and area
# (constantAREA)
# scalebar.txt contains length of scalebar in um on first line, and in
# pixels on the second line.
if os.path.isfile(folderIN + "/" + "scalebar.txt"):
   scale = []
   scaleBAR = open(folderIN + "/" + "scalebar.txt", 'r')
   for linear in scaleBAR:
      scale.append(float(linear))
   scaleBAR.close()
   constantLINEAR = (scale[0])/(scale[1]) # ratio between um and pixels
   constantAREA = constantLINEAR ** 2
constantLINEARstring = str(constantLINEAR)
constantAREAstring = str(constantAREA)

# Add fields and calculate protein area in arbitrary units and
# micrometers squared
arcpy.AddField_management(polySHP, "ProtAreaSc", "DOUBLE")
arcpy.CalculateField_management(polySHP, "ProtAreaSc",'!shape.area!',
    "PYTHON_9.3")
arcpy.AddField_management(polySHP, "ProtAreaMi", "DOUBLE")
arcpy.CalculateField_management(polySHP, "ProtAreaMi",'!shape.area!*'
    + constantAREAstring, "PYTHON_9.3")
arcpy.AddField_management(polySHP, "areaProt", "DOUBLE")
arcpy.CalculateField_management(polySHP, "areaProt",'!shape.area!*' +
    constantAREAstring, "PYTHON_9.3")
arcpy.FeatureToPoint_management(polySHP, folderIN + "/" +
    "centroids.shp")

# Calculating euclidean distances
featureTOlineTEMP = folderIN + "/" + "ftl.shp"
arcpy.FeatureToLine_management(shpBORDER, featureTOlineTEMP)
outEucDistance = EucDistance(featureTOlineTEMP, "", grainTIF)
```

```
outExtractByMask = ExtractByMask(outEucDistance, shpBORDER)
outZonalTableTEMP = folderIN + "/" + "ozt.dbf"
ZonalStatisticsAsTable(shpBORDER, "FID", outExtractByMask,
    outZonalTableTEMP, "DATA", "ALL")
distRows = arcpy.da.SearchCursor(outZonalTableTEMP, ['RANGE'])
distRow = distRows.next()
maximumDIST = distRow[0] # Maximum width of grain, aleurone to aleurone

dist = []
if numbZONES == 0:
   # Reading zone distances text file, if present
   txtFILE = open(txtFILEdistances, 'r')
   for line in txtFILE:
      dist.append(float(line)*(-1))
   txtFILE.close()
else:
   # Automatically calculating zones based on maximumDIST
   dist.append(-0.001)
   zoneWIDTH = maximumDIST/(float(numbZONES))
   for z in range(1,numbZONES):
      dist.append((-1)*z*zoneWIDTH)

zoneSHP = folderIN + "/" + "zones.shp"
arcpy.MultipleRingBuffer_analysis(shpBORDER, zoneSHP, dist,"", "",
    "ALL")
protein_in_ZONES = folderIN + "/" + "proteinZONES.shp"
arcpy.Intersect_analysis([zoneSHP,polySHP], protein_in_ZONES)
arcpy.CalculateField_management (protein_in_ZONES, "areaProt",
    '!shape.area!*' + constantAREAstring, "PYTHON_9.3")

# Add fields and calculate zone distance in arbitrary units and
# micrometers
arcpy.AddField_management(zoneSHP, "distZoneSc", "DOUBLE", 18,
    10,"","","NULLABLE")
arcpy.CalculateField_management (zoneSHP, "distZoneSc", '!distance! *
    (-1)', "PYTHON_9.3")
arcpy.AddField_management(zoneSHP, "distZoneMi", "DOUBLE", 18,
    10,"","","NULLABLE")
arcpy.CalculateField_management (zoneSHP, "distZoneMi", '!distance!*
    (-1)*' + constantLINEARstring, "PYTHON_9.3")
```

```
# Add fields and calculate zone area in arbitrary units and micrometers
arcpy.AddField_management(zoneSHP, "zoneAreaSc", "DOUBLE", 18,
    10,"","","NULLABLE")
arcpy.CalculateField_management (zoneSHP, "zoneAreaSc",
    '!shape.area!', "PYTHON_9.3")
arcpy.AddField_management(zoneSHP, "zoneAreaMi", "DOUBLE", 18,
    10,"","","NULLABLE")
arcpy.CalculateField_management (zoneSHP, "zoneAreaMi",
    '!shape.area!*' + constantAREAstring, "PYTHON_9.3")

# Add fields and calculate protein area in each zone in arbitrary units
# and micrometers
protein_in_zonesdissolved = folderIN + "/" + "proteinsZONES.shp"
arcpy.Dissolve_management(protein_in_ZONES, protein_in_zonesdissolved,
    "distance")
arcpy.AddField_management(protein_in_zonesdissolved, "protAreaSc",
    "DOUBLE", 18, 10,"","","NULLABLE")
arcpy.CalculateField_management (protein_in_zonesdissolved,
    "protAreaSc", '!shape.area!', "PYTHON_9.3")
arcpy.AddField_management(protein_in_zonesdissolved, "protAreaMi",
    "DOUBLE", 18, 10,"","","NULLABLE")
arcpy.CalculateField_management (protein_in_zonesdissolved,
    "protAreaMi", '!shape.area!*' + constantAREAstring, "PYTHON_9.3")

# Join zone background and protein areas and distances. Calculate
# percentage protein per zone
joinedRESULT = folderIN + "/" + "result.shp"
arcpy.SpatialJoin_analysis(zoneSHP, protein_in_zonesdissolved,
    joinedRESULT,"","","","CONTAINS")
arcpy.AddField_management(joinedRESULT, "percent", "DOUBLE", 18,
    10,"","","NULLABLE")
arcpy.CalculateField_management (joinedRESULT, "percent",
    '(!protAreaMi!/!zoneAreaMi!)*100', "PYTHON_9.3")
arcpy.DeleteField_management(joinedRESULT,["OID", "Join_Count",
    "TARGET_FID", "distance", "distance_1"])

# Output results
arcpy.TableToTable_conversion(joinedRESULT, folderIN, description +
    "_zones.csv")
```

```
# Clip centroids to outline of grain
arcpy.Clip_analysis(folderIN + "/" + "centroids.shp", shpBORDER,
    folderIN + "/" + "centroids2.shp")

# Load distance values for each protein body point
ExtractValuesToPoints(folderIN + "/" + "centroids2.shp",
    outEucDistance, folderIN + "/" + "proteinCentroids.shp")

# Add fields and calculate protein body distances and sizes, and
# description of treatment
arcpy.AddField_management(folderIN + "/" + "proteinCentroids.shp",
    "dist_Sc", "FLOAT",18, 10,"","","NULLABLE")
arcpy.AddField_management(folderIN + "/" + "proteinCentroids.shp",
    "dist_Mi", "FLOAT",18, 10,"","","NULLABLE")
arcpy.AddField_management(folderIN + "/" + "proteinCentroids.shp",
    "treatment", "TEXT","","","","","NULLABLE")
arcpy.CalculateField_management(folderIN + "/" +
    "proteinCentroids.shp", "dist_Sc","!RASTERVALU!","PYTHON_9.3")
arcpy.CalculateField_management(folderIN + "/" +
    "proteinCentroids.shp", "dist_Mi", "!RASTERVALU!*" +
    constantLINEARstring, "PYTHON_9.3")
arcpy.CalculateField_management(folderIN + "/" +
    "proteinCentroids.shp", "treatment", "'" + description + "'" ,
    "PYTHON_9.3")
arcpy.DeleteField_management(folderIN + "/" + "proteinCentroids.shp",
    ["ID", "GRIDCODE", "ORIG_FID", "RASTERVALU", "areaProt"])

# Output results
arcpy.TableToTable_conversion(folderIN + "/" + "proteinCentroids.shp",
    folderIN, description + "_spatial.csv")

arcpy.SpatialJoin_analysis(folderIN + "/" + "shape.shp", folderIN +
    "/" + "proteinCentroids.shp", folderIN + "/" + "shapeTemp.shp")
arcpy.Clip_analysis(folderIN + "/" + "shapeTemp.shp", shpBORDER,
    folderIN + "/" + "proteins.shp")
arcpy.DeleteField_management(folderIN + "/" +
    "proteins.shp",["TARGET_FID", "ID", "GRIDCODE", "Join_Count",
    "areaProt_1"])
```

```
        # Delete temporary files used
        arcpy.Delete_management(folderIN + "/" + "centroids.shp")
        arcpy.Delete_management(folderIN + "/" + "centroids2.shp")
        arcpy.Delete_management(folderIN + "/" + "ftl.shp")
        arcpy.Delete_management(folderIN + "/" + "ozt.dbf")
        arcpy.Delete_management(protein_in_ZONES)
        arcpy.Delete_management(folderIN + "/" + "shapeTemp.shp")
        arcpy.Delete_management(folderIN + "/" + "shape.shp")
        arcpy.Delete_management(joinedRESULT)

        arcpy.RefreshCatalog(folderIN)

        return

# RescalingBatch class used to rescale a batch of microscopy images to 1x1
# cell sizes
# This is required for input images used in the Protein class to ensure
# correct scaling from pixels to micrometers
class RescalingBatch(object):
    def __init__(self):
        self.label = 'Batch rescaling image to 1 x 1 pixel size'
        self.canRunInBackground = False

    def getParameterInfo(self):
        '''Set up the parameters and return the list of parameter objects.'''

        # Input grain raster
        param0 = arcpy.Parameter()
        param0.name = 'tifFile'
        param0.displayName = 'Input grain TIF file:'
        param0.parameterType = 'Required'
        param0.direction = 'Input'
        param0.datatype = 'DEFolder'
        param0.filter.list = ['tif', 'tiff', 'TIF']

        return [param0]

    def isLicensed(self):
        """Prevent the tool from running if the Spatial Analyst extension is
           not available."""
```

```
        if arcpy.CheckExtension('Spatial') == 'Available':
            return True # The tool can be executed.
        else:
            return False # The tool can not be executed.

    def updateParameters(self, parameters):
        return

    def updateMessages(self, parameters):
        return

    def execute(self, parameters, messages):
      # Calculations
      folder = parameters[0].valueAsText
      arcpy.env.cellSize = 1

      for input_file in glob.glob(os.path.join(folder,'*.tif')):
         filename = os.path.basename(input_file)
         input = input_file
         a = Raster(input)
         rescaleRATIO = 1/a.meanCellHeight
         b = folder + "/" + "pxl_1x1_" + filename
         arcpy.Rescale_management(a,b,str(rescaleRATIO),str(rescaleRATIO))

      return
```

Appendix D: Spatial analysis grain width calculator

Python script to output a list of the maximum grain width associated with each treatment. The resulting CSV file is then used by the CSV joining applications presented in appendices E and F to calculate the width of the five endosperm zones which are used in the presentation of the results.

```
# Script to calculate and store maximum zone width for calculation of zone
# width
# Takes arguments (1) input directory, (2) output file

import csv
import glob
import os
import sys

# Define inputs
input_path = sys.argv[1]
output_file = sys.argv[2]

# Remove trailing slashes from input_path
input_path_length = len(input_path) if input_path [-1] == "/" \
or input_path [-1] == "\\" else len(input_path) + 1

# Loop through input files
for input_file in glob.glob(os.path.join(input_path,'*.csv')):
   with open(input_file,'rU') as csv_file:
      filereader = csv.reader(csv_file)
      next(filereader)
      # Read maxWidth value
      maxWidth = max(float(column[4].replace(',', '')) for column in
          filereader)
      name, ext = os.path.splitext(input_file)
      # Get treatment name from filename
      treatment = "{:<12}".format(name [+input_path_length:])
      header = next(filereader,None)
      # Print treatment name and maxWidth to file
      with open(output_file, 'ab') as csvfile:
         filewriter = csv.writer(csvfile)
```

```
        filewriter.writerow([treatment, maxWidth])
```

Appendix E: Spatial analysis CSV file joiner script (1)

Python script for combining the individual CSV files produced by the protein body size-distribution analysis into a single CSV file for analysis. Script reads the treatment from the filename of each input file (in the format of ConN1T3R1i1), saves this information into a single file, and assigns each protein body measruement with a zone (used in figures only). The script takes inputs of a directory containing the CSV files to be joined, an input file containing the maximum grain width for each treatment combination (see appendix D), and outputs a single CSV file containing all data from the protein body size-distribution analysis.

```
# Script used to combine multiple protein body size-distribution CSV files
# from the controlled-environment experiment.
# Reads treatment from filename (e.g. ConN1T3R1i1.csv), separates, and
# appends to output CSV file.
# Assigns each measured protein body to a zone (1--5) based on input file
# of zone widths for each treatment.
# Takes arguments (1) input directory of CSV files to be joined, (2) input
#  file of zone wideths, (3) output file location.

import csv
import glob
import os
import sys

# Define inputs
input_path = sys.argv[1]
input_zones = sys.argv[2]
output_file = sys.argv[3]

# Remove trailing slashes from input_path
input_path_length = len(input_path) if input_path [-1] == "/" \
or input_path [-1] == "\\" else len(input_path) + 1

# Open output file
filewriter = csv.writer(open(output_file,'w', newline=''))
file_counter = 0

# Create dictionary of zone widths
zonedict = {}
```

```
with open(input_zones, 'rU') as csv_zones:
    zonereader = csv.reader(csv_zones)
    for row in zonereader:
        key, maxDist = row
        z1=float(maxDist)/5
        z2=2*(float(maxDist)/5)
        z3=3*(float(maxDist)/5)
        z4=4*(float(maxDist)/5)
        z5=float(maxDist)
        zonedict[key] = [float(z1), float(z2), float(z3), float(z4), float(z5)]

# Loop through input csv files
for input_file in glob.glob(os.path.join(input_path,'*.csv')):
    with open(input_file,'rU') as csv_file:
        filereader = csv.reader(csv_file)
        # Remove extension from input csv file
        name, ext = os.path.splitext(input_file)
        treatment = "{:<12}".format(name [+input_path_length:]).rstrip()
        namelong = "{:<12}".format(name [+input_path_length:])
        nameshort = name[+input_path_length:-1]
        # Access and assign treatment parameters from input filename
        temperature = namelong[:-9]
        nitrogen = namelong[+3:-7]
        timepoint = namelong[+5:-5]
        experimentalrep = namelong[+7:-3]
        imagerep = nameshort[+9:]
        analysisrep = name[-1]
        # Assign elements of experimental structure
        room = 1 if temperature == "Con" else 2
        pot = 0
        pot = 1 if nitrogen == "N1" and timepoint == "T2" else pot
        pot = 2 if nitrogen == "N2" and timepoint == "T2" else pot
        pot = 3 if nitrogen == "N1" and timepoint == "T3" else pot
        pot = 4 if nitrogen == "N2" and timepoint == "T3" else pot

        # Access and assign zone measurements for appropriate treatment
        z1 = float(zonedict[treatment][0])
        z2 = float(zonedict[treatment][1])
        z3 = float(zonedict[treatment][2])
        z4 = float(zonedict[treatment][3])
```

```
z5 = float(zonedict[treatment][4])

# Write data to output file
if file_counter < 1: # For first input file
   for i, row in enumerate(filereader):
      if i==0: # Write column headers to first row
         row.append('Temperature'),
         row.append('Nitrogen'),
         row.append('Timepoint'),
         row.append('ExperimentalRep'),
         row.append('ImageRep'),
         row.append('AnalysisRep'),
         row.append('Room'),
         row.append('Pot'),
         row.append('Zone')
      else: # Append treatment details to input data
         row.append(temperature),
         row.append(nitrogen),
         row.append(timepoint),
         row.append(experimentalrep),
         row.append(imagerep),
         row.append(analysisrep),
         row.append(room),
         row.append(pot),
         # Calculate and add zone of each protein body
         dist = float(row[5])
         if dist >=0 and dist < z1:
            row.append(1)
         elif dist >= z1 and dist < z2:
            row.append(2)
         elif dist >= z2 and dist < z3:
            row.append(3)
         elif dist >= z3 and dist < z4:
            row.append(4)
         elif dist >= z4:
            row.append(5)
      filewriter.writerow(row)
else: # For subsequent input files
   # Skip first row (column headers)
   header = next(filereader,None)
```

```
        for row in filereader:
            # Append treatment details to input data
            row.append(temperature),
            row.append(nitrogen),
            row.append(timepoint),
            row.append(experimentalrep),
            row.append(imagerep),
            row.append(analysisrep),
            row.append(room),
            row.append(pot),
            # Calculate and add zone of each protein body
            dist = float(row[5])
            if dist >=0 and dist < z1:
                row.append(1)
            elif dist >= z1 and dist < z2:
                row.append(2)
            elif dist >= z2 and dist < z3:
                row.append(3)
            elif dist >= z3 and dist < z4:
                row.append(4)
            elif dist >= z4:
                row.append(5)
            filewriter.writerow(row)
    file_counter += 1
```

Appendix F: Spatial analysis CSV file joiner script (2)

Python script functionally identical to the script presented in appendix E, but for the processing of CSV files from the WGIN diversity field trial experiment.

```
# Script used to combine multiple protein body size-distribution CSV files
# from the WGIN diversity field trial experiment.
# Reads treatment from filename (e.g. 15CaN2R1I1A1_spatial.csv), separates,
# and appends to output CSV file.
# Assigns each measured protein body to a zone (1--5) based on input file
# of zone widths for each treatment.
# Takes arguments (1) input directory of CSV files to be joined, (2) input
# file of zone widths, (3) output file location.

import csv
import glob
import os
import sys

# Define inputs
input_path = sys.argv[1]
input_zones = sys.argv[2]
output_file = sys.argv[3]

# Remove trailing slashes from input_path
input_path_length = len(input_path) if input_path [-1] == "/" \
or input_path [-1] == "\\" else len(input_path) + 1

# Open output file
filewriter = csv.writer(open(output_file,'w', newline=''))
file_counter = 0

# Create dictionary of zone widths
zonedict = {}
with open(input_zones, 'rU') as csv_zones:
   zonereader = csv.reader(csv_zones)
   for row in zonereader:
      key, maxDist = row
      z1=float(maxDist)/5
```

```
        z2=2*(float(maxDist)/5)
        z3=3*(float(maxDist)/5)
        z4=4*(float(maxDist)/5)
        z5=float(maxDist)
        zonedict[key] = [float(z1), float(z2), float(z3), float(z4), float(z5)]

# Loop through input csv files
for input_file in glob.glob(os.path.join(input_path,'*.csv')):
    with open(input_file,'rU') as csv_file:
        filereader = csv.reader(csv_file)
        # Remove extension from input csv file
        fullName, ext = os.path.splitext(input_file)
        # Remove _spatial appendix to filename
        name = fullName [+13:-8]
        namelong = "{:<12}".format(name [+input_path_length:])
        nameshort = name[+input_path_length:-1]
        # Access and assign treatment parameters from input filename
        year = name[:-10]
        genotype = name[+2:-8]
        nitrogen = name[+4:-6]
        block = name[+6:-4]
        imageRep = name[+8:-2]
        analysisRep = name[+10:]
        # Assign elements of experimental structure
        yearTreat = 1 if year == ''15'' else 2
        main = 1 if nitrogen == "N2" else 2
        split = 0
        split = 1 if genotype == "Ca" else split
        split = 2 if genotype == "He" else split
        split = 3 if genotype == "Is" else split
        split = 4 if genotype == "Ss" else split
        # Access and assign zone measurements for appropriate treatment
        z1 = float(zonedict[treatment][0])
        z2 = float(zonedict[treatment][1])
        z3 = float(zonedict[treatment][2])
        z4 = float(zonedict[treatment][3])
        z5 = float(zonedict[treatment][4])

        # Write data to output file
        if file_counter < 1: # For first input file
```

```
    for i, row in enumerate(filereader):
        if i==0: # Write column headers to first row
            row.append('Year')
            row.append('YearTreat')
            row.append('Genotype'),
            row.append('Nitrogen'),
            row.append('Timepoint'),
            row.append('Block'),
            row.append('ImageRep'),
            row.append('AnalysisRep'),
            row.append('Main'),
            row.append('Split'),
            row.append('Zone')
        else: # Append treatment details to input data
            row.append(year)
            row.append(yearTreat)
            row.append(genotype),
            row.append(nitrogen),
            row.append(timepoint),
            row.append(block),
            row.append(imageRep),
            row.append(analysisRep),
            row.append(main),
            row.append(split),
            # Calculate and add zone of each protein body
            dist = float(row[5])
            if dist >=0 and dist < z1:
                row.append(1)
            elif dist >= z1 and dist < z2:
                row.append(2)
            elif dist >= z2 and dist < z3:
                row.append(3)
            elif dist >= z3 and dist < z4:
                row.append(4)
            elif dist >= z4:
                row.append(5)
        filewriter.writerow(row)
    else: # For subsequent input files
        # Skip first row (column headers)
        header = next(filereader,None)
```

```
    for i, row in enumerate(filereader):
          # Append treatment details to input data
          row.append(year)
          row.append(yearTreat)
          row.append(genotype),
          row.append(nitrogen),
          row.append(timepoint),
          row.append(block),
          row.append(imageRep),
          row.append(analysisRep),
          row.append(main),
          row.append(split),
          # Calculate and add zone of each protein body
          dist = float(row[5])
          if dist >=0 and dist < z1:
             row.append(1)
          elif dist >= z1 and dist < z2:
             row.append(2)
          elif dist >= z2 and dist < z3:
             row.append(3)
          elif dist >= z3 and dist < z4:
             row.append(4)
          elif dist >= z4:
             row.append(5)
          filewriter.writerow(row)
file_counter += 1
```

Appendix G: Zone analysis CSV file joiner script

Since the image analysis software method was modified between the analysis of the controlled-environment and the WGIN diversity field experiment to provide automated output of the results of the total protein concentration gradient (by zone) analysis, a Python script was written to combine the multiple CSV files output by this analysis.

Python 3 script for combining the CSV files produced as from a single year of the WGIN diversity field trial experiment generated by the script presented in appendix F.

```
# Script used to combine multiple protein concentration distribution CSV
    files
# from the WGIN diversity field trial experiment

# Reads treatment from filename (e.g. 15CaN2R1I1A1_zones.csv), separates,
# and appends to output CSV file

# Takes arguments (1) input directory of CSV files to be joined, (2) output
    file
# location

import csv
import glob
import os
import sys

# Define inputs
input_path = sys.argv[1]
output_file = sys.argv[2]

# Remove trailing slashes from input_path
input_path_length = len(input_path) if input_path [-1] == "/" \
or input_path [-1] == "\\" else len(input_path) + 1

# Open output file
filewriter = csv.writer(open(output_file,'wb'))
file_counter = 0

# Loop through input csv files
for input_file in glob.glob(os.path.join(input_path,'*.csv')):
```

```
with open(input_file,'rU') as csv_file:
    filereader = csv.reader(csv_file)
    # Remove extension from input csv file
    fullName, ext = os.path.splitext(input_file)
    # Remove _zones appendix to filename
    name = fullName [+3:-6]
    # Access and assign treatment parameters from input filename
    year = name[:-10]
    genotype = name[+2:-8]
    nitrogen = name[+4:-6]
    block = name[+6:-4]
    imageRep = name[+8:-2]
    analysisRep = name[+10:]
    # Assign elements of experimental structure
    main = 1 if nitrogen == "N2" else 2
    split = 0
    split = 1 if genotype == "Ca" else split
    split = 2 if genotype == "He" else split
    split = 3 if genotype == "Is" else split
    split = 4 if genotype == "Ss" else split

    # Loop through zone distance measurements to calculate zone width
    for i, row in enumerate(filereader):
        if i == 0:
        # new file, ignore header row, reset zoneA and zoneB
            zoneA = 0
            zoneB = 0
        if i == 1:
        # read distance of last zone from aleurone layer
            zoneA = row[2]
        if i == 2:
        # read distance of last but one zone from aleurone layer and
            calculate
        # zoneWidth
            zoneB = row[2]
            zoneWidth = (float(zoneA) - float(zoneB))

    # Return to start of file
    csv_file.seek(0)
```

```
# Write data to output file
if file_counter < 1: # For first input file
   for i, row in enumerate(filereader):
      if i==0: # Write column headers to first row
         row.append('ZoneDist')
         row.append('Year')
         row.append('Genotype'),
         row.append('Nitrogen'),
         row.append('Timepoint'),
         row.append('Block'),
         row.append('ImageRep'),
         row.append('AnalysisRep'),
         row.append('Main'),
         row.append('Split')
      else:
         # Calculate zoneDist
         zoneDist = (0.5*zoneWidth) + ((5-i)*zoneWidth)
         # Append treatment details to input data
         row.append(zoneDist),
         row.append(year),
         row.append(genotype),
         row.append(nitrogen),
         row.append(timepoint),
         row.append(block),
         row.append(imageRep),
         row.append(analysisRep),
         row.append(main),
         row.append(split)
      filewriter.writerow(row)
else: # For subsequent input files
   # Skip first row (column headers)
   header = next(filereader,None)
   for i, row in enumerate(filereader):
      # Calculate zoneDist
      zoneDist = (0.5*zoneWidth) + ((4-i)*zoneWidth)
      # Append treatment details to input data
      row.append(zoneDist),
      row.append(year),
      row.append(genotype),
      row.append(nitrogen),
```

```
            row.append(timepoint),
            row.append(block),
            row.append(imageRep),
            row.append(analysisRep),
            row.append(main),
            row.append(split)
            filewriter.writerow(row)
    file_counter += 1
```

Appendix H: Journal of Experimental Botany publication

Journal of Experimental Botany, Vol. 69, No. 12 pp. 3117–3126, 2018
doi:10.1093/jxb/ery127 Advance Access publication 5 April 2018
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

RESEARCH PAPER

Temperature and nitrogen supply interact to determine protein distribution gradients in the wheat grain endosperm

George P. Savill[1], Adam Michalski[2], Stephen J. Powers[3], Yongfang Wan[1], Paola Tosi[4], Peter Buchner[1] and Malcolm J. Hawkesford[1,*]

[1] Plant Sciences Department, Rothamsted Research, West Common, Harpenden, AL5 2JQ, UK
[2] Institute of Geodesy and Geoinformatics, Wrocław University of Environmental and Life Sciences, ul. Grunwaldzka 53, 50–357 Wrocław, Poland
[3] Computational and Analytical Sciences Department, Rothamsted Research, West Common, Harpenden, AL5 2JQ, UK
[4] School of Agriculture, Policy, and Development, University of Reading, Earley, Reading, RG6 6BZ, UK

* Correspondence: malcolm.hawkesford@rothamsted.ac.uk

Received 29 November 2017; Editorial decision 21 March 2018; Accepted 21 March 2018

Editor: Greg Rebetzke, CSIRO Agriculture and Food, Australia

Abstract

Gradients exist in the distribution of storage proteins in the wheat (*Triticum aestivum*) endosperm and determine the milling properties and protein recovery rate of the grain. A novel image analysis technique was developed to quantify both the gradients in protein concentration, and the size distribution of protein bodies within the endosperm of wheat plants grown under two different (20 or 28 °C) post-anthesis temperatures, and supplied with a nutrient solution with either high or low nitrogen content. Under all treatment combinations, protein concentration was greater in the endosperm cells closest to the aleurone layer and decreased towards the centre of the two lobes of the grain, i.e. a negative gradient. This was accompanied by a decrease in size of protein bodies from the outer to the inner endosperm layers in all but one of the treatments. Elevated post-anthesis temperature had the effect of increasing the magnitude of the negative gradients in both protein concentration and protein body size, whilst limiting nitrogen supply decreased the gradients.

Keywords: Breadmaking quality, gluten protein, grain protein distribution, protein body size distribution, protein distribution gradient, *Triticum aestivum*, wheat endosperm, wheat grain protein.

Introduction

The starchy endosperm is the central storage tissue of the wheat grain and is the material from which white flour is produced. Gradients exist in the distribution of gluten storage proteins within the starchy endosperm of developing and mature grain: these gradients are both qualitative and quantitative in nature (Tosi *et al.*, 2011; He *et al.*, 2013), and influence both the milling properties and milling yield of the grain. The milling of wheat grain is achieved through a sequence of milling and sieving stages, which results in the production of numerous mills streams. These individual mill streams are enriched with different parts of the wheat grain, and so gradients in endosperm protein concentration and composition will result in the production of mill streams with specific protein content, composition, and baking characteristics (Yahata *et al.*, 2006; Wang *et al.*, 2007; Wan *et al.*, 2014; Zhou *et al.*, 2018). Different mill streams are then blended

by millers to produce flour with the desired qualities, and so understanding the factors affecting the composition of mill streams is of great importance. Furthermore, during milling a proportion of the endosperm tissue remains adhered to the bran layers, and since protein is concentrated in the outer layers of the endosperm (Kent, 1966; Ugalde and Jenner, 1990*a*; Tosi *et al.*, 2011; He *et al.*, 2013), a proportionally greater amount of protein relative to starch is lost during the production of white flour. Hence, any increase in the gradient of protein concentration may result in an even greater amount of protein being lost during milling.

Whilst the presence of a gradient in protein concentration within the wheat grain has been known for some time (Cobb, 1905), the mechanisms behind this gradient remain unexplained. Recently Moore *et al.* (2016) used ^{15}N labelling with NanoSIMS analysis to demonstrate that glutamine, the most abundant form of nitrogen within the wheat plant (Fisher and Macnicol, 1986), is transported radially from the endosperm cavity, across the starchy endosperm, before becoming concentrated in the subaleurone cells. This study, which focussed on limited microscale transects, demonstrated that the subaleurone cells had larger protein bodies, and it was speculated that they acted as a strong sink driving amino acid transport across the endosperm.

The future of wheat production for bread-making is likely to face many challenges, such as limits on the use of nitrogen fertiliser and changing climate, which, in the UK, will include increases in ambient temperatures experienced during grain filling. The production of nitrogen fertiliser is intensive in its energy requirement, and at present relies heavily on the use of fossil fuels (Dawson and Hilton, 2011). Due to this, both financial and policy constraints are likely to limit the amount of nitrogen fertiliser that can be applied to future bread-making wheat crops. Additionally, it is predicted that there will be a global increase in temperature (Alexander *et al.*, 2006) and an increase in temperature variability, with more frequent extreme weather events such as heatwaves and droughts (Hennessy *et al.*, 2008); these changes to our climate are likely to be to the detriment of agriculture and food security (Wheeler and von Braun, 2013)

Elevated temperature affects the development of wheat, drastically shortening the grain-filling period, and resulting in altered grain expression and related protein accumulation (Altenbach, 2012). Nitrogen nutrition has been shown to affect the deposition of protein within the grain, with higher levels of nitrogen supply resulting in an increase in ω-gliadin accumulation in the outer layers of the starchy endosperm (Wan *et al.*, 2014). The effect of temperature on the gradients of gluten storage proteins within the wheat endosperm has not been previously described.

Previous analysis of protein gradients in wheat grain has been a low-throughput process, reliant on micro-dissection (Cobb, 1905; Ugalde and Jenner, 1990*a*, 1990*b*), sub-sampling of microscopy images (Tosi *et al.*, 2011), or pearl milling (He *et al.*, 2013). Here, a high-throughput image analysis technique for quantifying features in light microscopy sections is presented, which analyses both the gradient in protein concentration and the size distribution of protein bodies within the endosperm of wheat grain during grain-filling. Wheat grain grown under controlled-environment conditions is used to describe the effect of elevated post-anthesis temperature and reduced nitrogen input on storage protein accumulation within the wheat endosperm in the context of UK climate and agronomical practice.

Materials and methods

Plant material and growth conditions

A spring wheat cultivar (*Triticum aestivum* L., 'Cadenza') was grown under controlled-climate conditions at Rothamsted Research, Harpenden, UK. A complete randomised block design was used to grow 480 plants in 96 pots split equally between three experimental blocks, with a treatment structure of two levels of nitrogen fertiliser crossed with two post-anthesis temperature regimes.

Plants were grown to anthesis in a single Weiss Gallenkamp controlled-environment room with a floor area of 16 m^2 and a height of 3 m. Day/night temperatures were set at 20/15 °C and humidity at 65/75%. The photoperiod was maintained at 16 h per day, and the light intensity was 500 µmol m^{-2} s^{-1} from 400 W HQI metal halide lamps with a supplementary 10% mix of tungsten lighting. The change between day and night temperatures was programmed to take 2 h, to approximate field conditions. At anthesis, which was determined when three out of five plants in a pot showed externally visible anthers, half of the plants were moved to a second, identical controlled-environment room and grown at day/night temperatures of 28/15 °C. This temperature regimen was chosen since it represents the minimum threshold temperatures for heatwaves in the UK (Public Health England, 2014). The remaining plants were kept in the original room at the control day/night temperature of 20/15 °C. Humidity and photoperiod remained unchanged between treatments, and the layout of pots within the room was maintained throughout.

Plants were grown in 25-cm pots with the nutrient-poor 'Rothamsted nematode mix' (80% sterilised loam, 15% sand, and 5% 5-mm grit), supplemented with either a high- or low-nitrogen liquid nutrient solution. Nutrient solution was applied from 10 d after sowing, at a rate of 500 ml once a week until ear emergence, and 500 ml twice a week from ear emergence to anthesis. Nine applications (4.5 l in total) of nutrient solution were made per pot of five plants, supplying 504 mg and 50.4 mg of nitrogen to the high- and low-nitrogen treatments respectively. The composition of the nutrient solutions was 4 mM $Ca(NO_3)_2.4H_2O$ (high-nitrogen treatment only), 0.4 mM $Ca(NO_3)_2.4H_2O$ (low-nitrogen treatment only), 3.6 mM $CaCl_2$ (low-nitrogen treatment only), 0.25 mM KH_2PO_4, 0.5 mM KOH, 0.75 mM $MgSO_4.7H_2O$, 0.03 mM $CaCl_2$, 0.1 mM FeNaEDTA, 30 µM H_3BO_3, 10 µM $MnSO_4.4H_2O$, 1 µM $ZnCl_2.7H_2O$, 3 µM $CuSO_4.5H_2O$, and 0.5 µM $Na_2MoO_2.2H_2O$. Pots were placed on saucers to prevent loss of nutrient solution, and plants were watered with deionised water as required to prevent the soil from drying out.

Sampling protocol

Grain was sampled for light-microscopy analysis from the first floret in spikelets from the centre of the ear, to ensure grain of comparable developmental status were selected for analysis. Samples were taken from the first ear to reach anthesis on each plant. One developing grain per ear and treatment combination per block (three replicates) was harvested at 14 and 21 d post-anthesis for the low-temperature (20 °C) treatment, and at 10 and 15 d post-anthesis for the high-temperature (28 °C) treatment. Sampling timepoints were adjusted to account for the difference in accumulated thermal time experienced by the plants in each treatment. Accumulated thermal time was calculated using a base temperature of 4.1 °C (Fischer, 1985), and rounded to the nearest day to determine sampling timepoints. Sixteen grains were also sampled for total nitrogen content analysis at the same timepoints used for microscopy analysis.

Morphological and nitrogen content measurements

At maturity, all five plants within a pot were harvested, hand-threshed, and dried at 80 °C to 4–5% moisture content. Grain from each pot was combined and used to measure thousand grain weight, grain yield, and final nitrogen content. A total grain count was calculated from the grain yield and thousand grain weight.

Nitrogen content of the wholemeal flour was determined by the Dumas method using a LECO CN628 Combustion Analyser (LECO Corporation, St Joseph, Michigan, USA), and is expressed as percent dry matter.

Microscopy

Developing grain sampled at two timepoints during grain development were fixed in 4% paraformaldehyde, 2.5% glutaraldehyde fixative, dehydrated in a graded ethanol series and embedded in medium-grade LR White Resin as described by Tosi *et al.* (2011). Sections of 1 µm thickness were cut and stained for protein with 1% Naphthol Blue Black in 7% (v/v) acetic acid for 30 s. Images were captured at 20× magnification on a Zeiss Axiophot light microscope (Zeiss, Oberkochen, Germany), and stitched into high-resolution composite images using MetaMorph Microscopy Automation and Image Analysis Software (Molecular Devices, Sunnyvale, CA, USA) and Adobe Photoshop (Adobe Systems Inc., San Jose, CA, USA). Composite images of the entire endosperm were then analysed.

Image analysis

Light-microscopy images of developing wheat grain stained for protein were analysed using a custom toolbox in ArcMap™, part of the ArcGIS™ 10.4 software package (ESRI®, Redlands, CA, USA) (available as an ArcGIS™ python toolbox, doi: 10.5281/zenodo.1066914). Images were loaded (Fig. 1A), and an outline was manually drawn around the endosperm of the grain, just within the aleurone layer (Fig. 1B). Two methods of image analysis were then run simultaneously: protein concentration distribution analysis, and protein body size distribution analysis. Protein concentration distribution analysis describes the overall gradients in protein concentration, whilst protein body size distribution analysis measures the area of each individual protein body, and records the distance of that protein body from the outline drawn around the endosperm.

Supervised maximum likelihood classification was used to detect protein within the light-microscopy images on a pixel-by-pixel basis (Fig. 1D). For this classification, multiple training samples were taken from each RGB image to give examples of both stained protein, and non-protein areas within the endosperm. These training samples were taken from 10 areas of protein across different parts of the grain to account for variation in stain intensity across the section, and from sufficient non-protein areas of the grain to account for all combinations of tone and intensity that did not represent protein. To account for variation in the defined training samples used for the image classification, each image was analysed three times using different training samples each time. Furthermore, every analysis instance was checked to verify the accuracy of the image classification in discriminating protein from background.

To describe the gradient in protein within the endosperm, protein concentration distribution analysis was used to measure and compare protein concentration in five zones within the endosperm. Five concentric bands were automatically rendered inwards from the endosperm outline to create five zones of equal width covering the range of outer to inner endosperm (Fig. 1C). These zones were then overlaid on the extracted protein body area data as determined by maximum likelihood classification (Fig. 1D), and protein within each zone was measured. The result is a measurement of protein concentration on a by-area basis within each of the five zones in the endosperm (Fig. 1E).

The use of five zones was determined to be optimal through experimentation with different numbers of zones (from three to seven) to find the number that was geometrically optimal to describe the gradient in protein concentration via best-fit regression modelling of total protein body area per zone versus zone number (data not shown).

For protein body size distribution analysis, the area of each protein body was measured, the midpoint detected (Fig. 1F), and the shortest Euclidean distance between this point and the outline of the endosperm calculated.

Data processing

Post-analysis, a conversion was applied to the data collected from the protein concentration gradient analysis: nitrogen content measurements made on grain sampled from the same plant and at the same timepoint were used to calculate a conversion factor between actual grain protein content and protein content as determined through microscopy-image analysis. A unique conversion factor was calculated for each combination of treatment and experimental replicate. This conversion factor was applied to all collected data to account for any differences in protein detection efficiency between treatments, since any changes in grain water content or protein density within a protein body would result in inaccuracies in the measurement of protein concentration on a by-area basis. The equation used for the conversion was as follows:

$$Conversion\ factor = \frac{Protein\ concentration\ from\ nitrogen\ content\ measurements}{100\times\left(\frac{Total\ protein\ area\ from\ microscopy\ image\ analysis}{Total\ grain\ area\ from\ microscopy\ image\ analysis}\right)}$$

To account for differences in the size of grain from different treatments, prior to analysis the zone attributed to each protein concentration measurement was converted from a factor to a variable by calculating the distance between the outline of the endosperm and the midpoint between the boundaries of each zone. This value was then converted from pixels to micrometres to give a meaningful mean physical distance for each of the five protein concentration measurements made on each image.

Statistical analysis

The GenStat® statistical software (2015, Eighteenth Edition, VSN International Ltd, Hemel Hempstead, UK) was used to analyse the effect of post-anthesis temperature and nitrogen fertiliser level on the total grain count, nitrogen yield, thousand grain weight, and total grain yield by ANOVA. After inspection of *F*-tests for the main effects and interactions between the treatment factors, the least-significant difference at the 5% (P=0.05) level calculated from the standard error of the difference between means on the residual degrees of freedom from the ANOVA was used to make comparisons of relevant means.

GenStat® was used to fit a linear mixed model to both the distribution of protein concentration over the calculated zones, and to the protein body size distribution data using the method of restricted maximum likelihood. Data from each image analysis technique were analysed separately.

In the analysis of the protein concentration distribution data, the calculated mean distance from the aleurone cell layer within the wheat grain was used rather than the zone factor. The fixed part of the linear mixed model was *Temperature*Nitrogen*Timepoint*Distance*, where *Temperature*, *Nitrogen*, and *Timepoint* are factors, *Distance* is a variable calculated from zone, and asterisks represent the main effects and interactions between the model terms. In the analysis of the protein body size distribution data, the protein body area measurements were log-transformed to account for some heterogeneity of variance over the treatment combinations. The fixed part of the linear mixed model was *Temperature*Nitrogen*Timepoint*ProteinDistance*, where *Temperature*, *Nitrogen*, and *Timepoint* are factors, and *ProteinDistance* is a variable. The random part of each model was (*Room.Block*)/*Pot*/*ImageRep*/*Zone*/*AnalysisRep* and (*Room.Block*)/*Pot*/*ImageRep*/*AnalysisRep* for protein concentration and body size, respectively, where *Room.Block* is the interaction between growth rooms and experimental blocks to account for the two rooms being used, *Pot* is the factor for pots within room by block combinations, *ImageRep* is

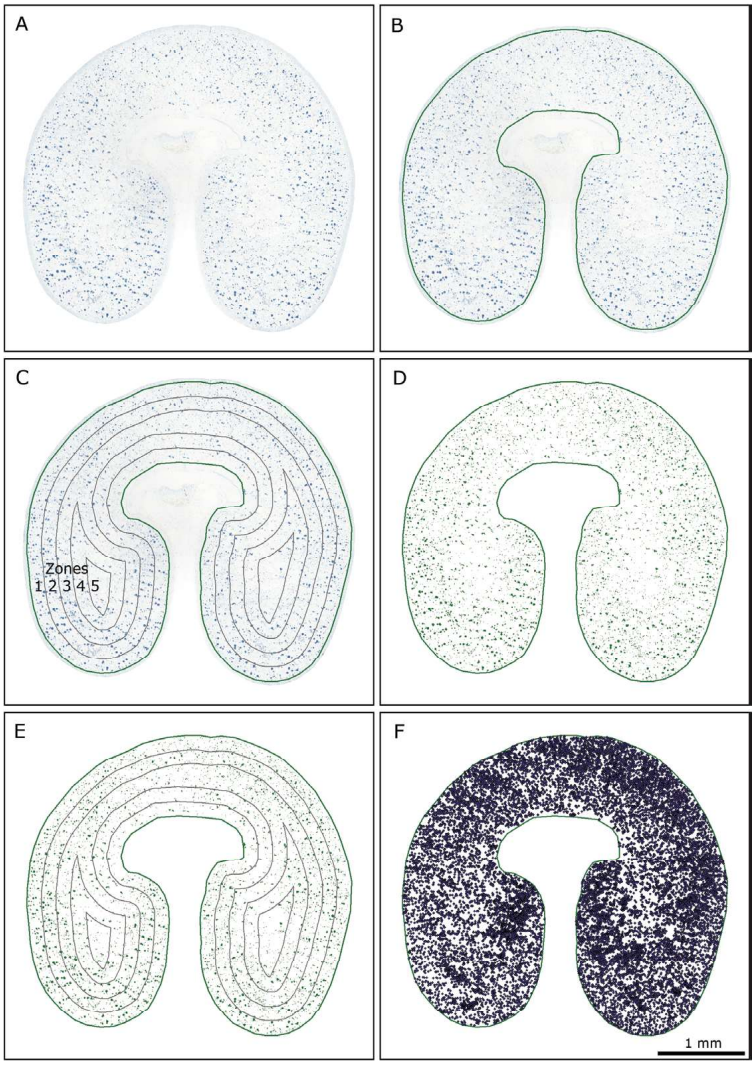


Fig. 1. Overview of the grain protein gradient analysis procedure. (A) A composite light-microscopy image of a 1-μm thick wheat grain section selectively stained for protein with 1% Naphthol Blue Black in 7% acetic acid (v/v) is loaded. (B) An outline is manually drawn around the endosperm, just inside the aleurone layer. (C) Five concentric bands are automatically drawn inwards from the outline of the endosperm, creating zones representing tissue from outer to inner endosperm (labelled 1–5). (D) Pixels representing protein are extracted using image classification, and the background is removed. (E) The zones from stage (C) are overlaid on the extracted protein from stage (D) and protein concentration values are calculated on an area basis for each zone within the endosperm. (F) For protein body size distribution analysis, individual protein bodies are detected, and their area and distance from the aleurone layer are measured. The example shown is a section from a grain under the high-nitrogen, 20 /15 °C temperature treatment, sampled at 21 d post-anthesis.

the technical replication of microscopy sections, *Zone* is the concentric zone within the endosperm of each image from which a measurement was taken, *AnalysisRep* is the technical replication of each image analysis, and slash (/) indicates nesting of these factors. Non-linearity with respect to distance of measurement from the aleurone was tested by adding a distance-squared term to the fixed part of the model, and no evidence of non-linearity was found ($P<0.05$, F-tests). Inspection of residual plots revealed that log-transformation of the protein body area data was adequate to stabilise their variance, and that no transformation was required for any of the other analyses.

Results

Measurements taken at harvest

Grain yield was reduced by both increased post-anthesis temperature and reduced nitrogen supply, with a significant interaction between these two factors ($P=0.008$, F-test): reducing the nitrogen application rate had a greater effect on yield in plants grown under 20 /15 °C post-anthesis temperature

conditions than in plants grown under 28 /15 °C post-anthesis temperature (Fig. 2A). Yield was higher under the control temperature treatment, regardless of nitrogen input level.

Thousand grain weight was reduced by an average of 11.97 g, from 40.06 g to 28.09 g, under high post-anthesis temperature (P<0.001, F-test) (Fig. 2C) and by 2.01 g, from 35.08 g to 33.07 g, under the low-nitrogen treatment (P=0.019, F-test) (Fig. 2D). No significant interaction between temperature and nitrogen treatment was found (P=0.170, F-test).

Mean grain count per plant was calculated, and was found to be reduced by approximately 10%, from 304 to 276 grains per plant, under the low-nitrogen treatment (P=0.002, F-test) (Fig. 2F); elevated post-anthesis temperature did not affect grain count (P=0.198, F-test) (Fig. 2E), and there was no interaction between nitrogen and temperature treatments (P=0.723, F-test).

Nitrogen concentration at maturity was higher in grain grown under elevated post-anthesis temperature, and in grain grown under higher levels of applied nitrogen, with a significant interaction between these two factors (P=0.032, F-test): increasing nitrogen application rate increased grain nitrogen concentration to a greater extent in plants grown under high post-anthesis temperature (from 2.23% to 2.56%) than in plants grown under control conditions (from 1.49% to 1.63%) (Fig. 2B). Grain nitrogen yield per plant was calculated from total grain yield and nitrogen concentration measurements, and was reduced by the low-nitrogen treatment (P<0.001, F-test) (Fig. 2H), but was not significantly different between post-anthesis temperature treatments (P=0.626, F-test) (Fig. 2G), and there was no significant interaction between nitrogen and temperature treatments (P=0.649, F-test).

Time to maturity was greatly reduced by an increased temperature post-anthesis, with an average time from anthesis to harvest at maturity of 46 d for the control temperature treatment, and 33 d for the elevated temperature treatment.

Grain protein concentration gradients

A significant four-way interaction was found between post-anthesis temperature, nitrogen treatment, sampling timepoint, and mean distance of protein concentration from the aleurone layer (P=0.049, F-test). Under all treatments, protein concentration was greatest nearest to the aleurone layer, and lowest farthest from the aleurone layer. This negative gradient in protein concentration was greater, with more protein located closer to the aleurone layer, both in grain grown under high temperatures, and in grain grown under high nitrogen. The effect of nitrogen supply on the gradient was greater in grain grown under higher temperatures. Furthermore, the negative gradient in protein concentration was greater, and the effect of both elevated post-anthesis temperature and nitrogen supply more pronounced at the later sampling timepoint. The protein concentration gradient results for low- and high-nitrogen treatments at the first sampling timepoint are shown in Fig. 3A and 3B, respectively, and Fig. 3C and 3D show the same for the second, later sampling timepoint.

Protein body size distribution

A significant four-way interaction was found between post-anthesis temperature treatment, nitrogen treatment, sampling timepoint, and distance of protein bodies from the aleurone layer (P<0.001, F-test). Under all treatment combinations, a gradient in protein body size was observed, with mean size increasing towards the aleurone layer in every treatment combination except for the control temperature, low-nitrogen,

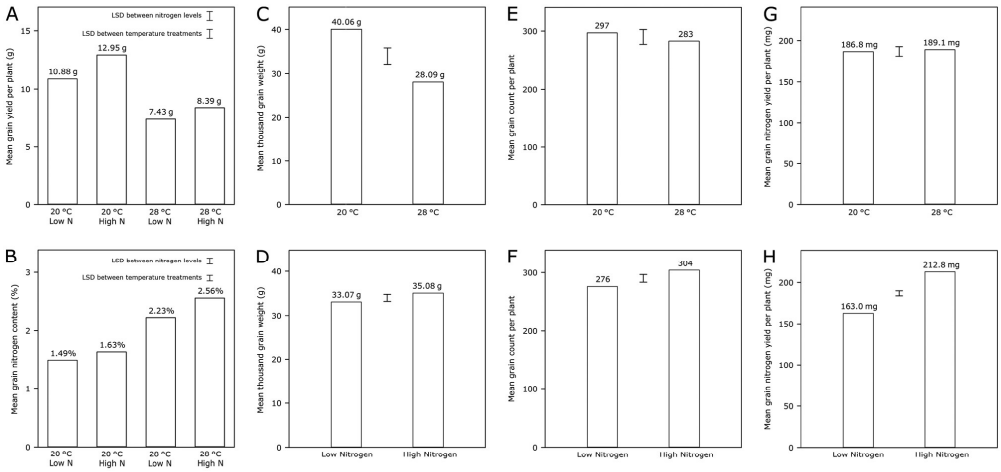


Fig. 2. Mature grain characteristics. (A) Mean grain yield (n=3) per plant for all combinations of temperature and nitrogen treatments. (B) Mean grain nitrogen content (n=3) for all combinations of temperature and nitrogen treatments. (C) Mean thousand grain weights (n=6) for the two temperature treatments, and (D) for the two nitrogen treatments (n=6). (E) Calculated mean total grain counts (n=6) per plant for the two temperature treatments, and (F) for the two nitrogen treatments (n=6). (G) Calculated mean grain nitrogen yield (n=6) per plant for the two temperature treatments, and (H) for the two nitrogen treatments (n=6). Bars for least-significant differences (P=0.05) are shown.

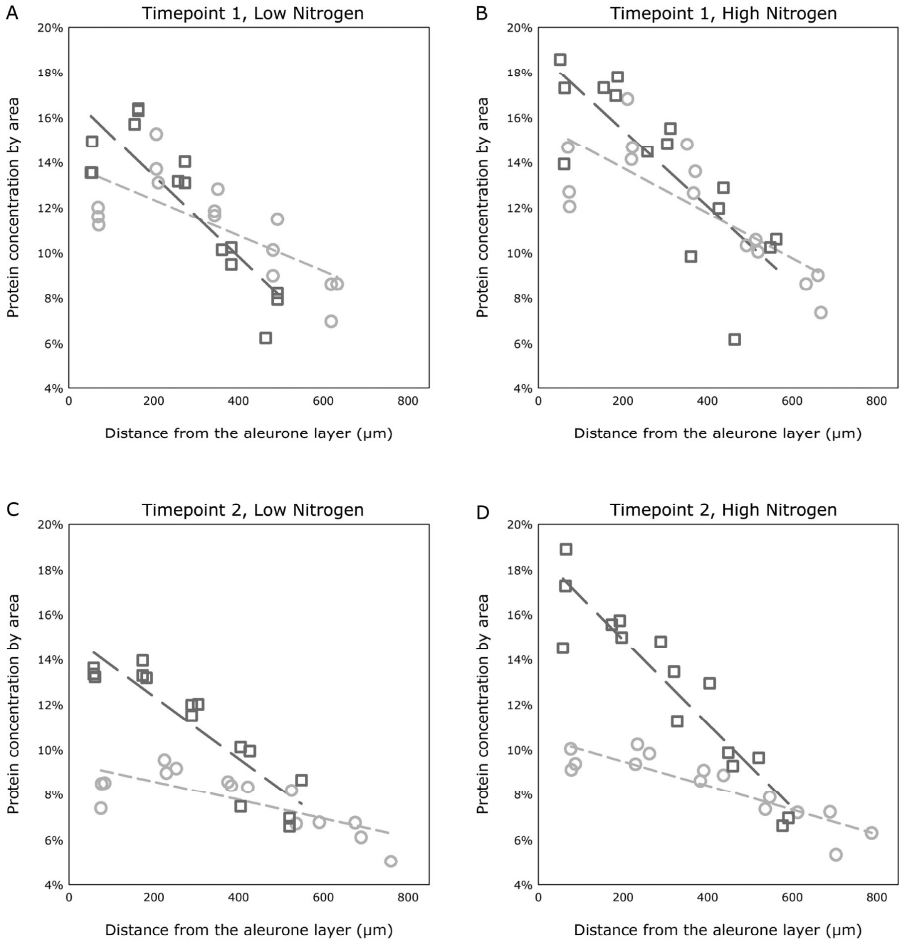


Fig. 3. Results of protein concentration gradient analysis. Grain sampled at timepoint 1 with low-nitrogen treatment (A) and high-nitrogen treatment (B), and at timepoint 2 with low-nitrogen treatment (C) and high-nitrogen treatment (D). Dark grey squares and long-dashed lines represent the 28 °C treatment, and light grey circles and short-dashed lines represent the control 20 °C treatment. The graphs in (A–D) show results of the same model, with portions of the results being separated for clarity. Trendlines represent the result of the linear mixed modelling analysis (using restricted maximum likelihood), which was applied to the entire dataset, whilst data points represent mean protein concentration values within each zone from each of the three experimental replicates analysed. Underlying biological variation for entire dataset, s^2=0.178.

later sampling timepoint treatment combination, where protein bodies showed a marginal decrease in size towards the aleurone layer. Protein bodies were larger at the later sampling timepoint for all treatments, and were larger in grain grown under control temperature at the earlier timepoint, but larger in grain grown at high temperature at the later timepoint. At the later sampling timepoint the effects of both post-anthesis temperature and nitrogen supply on protein body size distribution were more pronounced. Across both temperature treatments, grain provided with a higher supply of nitrogen showed greater gradients in the size distribution of protein bodies, with larger protein bodies closer to the aleurone layer, and smaller protein bodies towards the centre of the endosperm.

The results of the protein body size distribution analysis are presented in Fig. 4A and 4B for the low- and high-nitrogen treatments, respectively at the first sampling timepoint, and in Fig. 4C and 4D for the later sampling timepoint. The plotted means for grain sampled at the earlier timepoint (Fig.4A, B) show an initial decrease, followed by some evidence of an increase in mean protein body size as the distance from the aleurone layer increases beyond ~500 µm. This effect is only observed at the early sampling timepoint, with the plotted means from the later timepoint showing a more linear relationship with distance from the aleurone layer. These results are explored further in the histograms of protein body size distribution across the five endosperm zones shown in Fig. 5, which show how size is affected by each combination

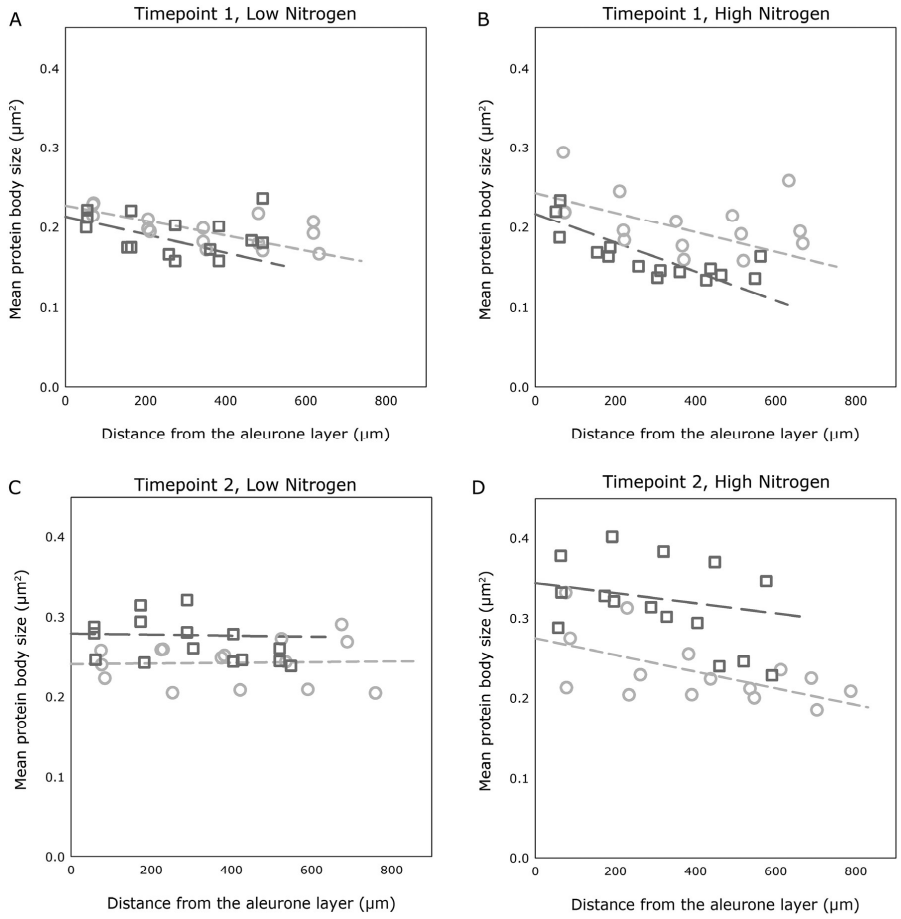


Fig. 4. Results of protein body size distribution gradient analysis. Grain sampled at timepoint 1 with low-nitrogen treatment (A) and high-nitrogen treatment (B), and at timepoint 2 with low-nitrogen treatment (C) and high-nitrogen treatment (D). Dark grey squares and long-dashed lines represent the 28 °C treatment, and light grey circles and short-dashed lines represent the control 20 °C treatment. The graphs in (A–D) show results of the same model, with portions of the results being separated for clarity. Trendlines represent the result of the linear mixed modelling analysis (using restricted maximum likelihood), which was applied to the entire dataset. This analysis found that separate linear trends was the best statistical representation of the temperature by nitrogen by timepoint interactions. Data points represent mean protein body size from each of the three experimental replicates within each zone. The analysis was completed on log-transformed data, and back-transformed data is presented for clarity. Underlying biological variation for entire dataset, s^2=2.345.

of temperature and nitrogen treatment at both the early and late sampling timepoint.

Discussion

Grain morphology

Elevated post-anthesis temperature produced the greatest change in grain morphology, with grain subjected to this treatment showing a greatly reduced total yield and thousand grain weight (Fig. 2A, C). These results indicate that the size of the mature grain was greatly reduced by increased temperatures during grain-filling. Limiting nitrogen supply prior to anthesis also reduced the grain size (Fig. 2D), albeit to a lesser degree than elevated temperature, and the effect of nitrogen input on grain yield was less apparent in plants grown under elevated post-anthesis temperatures (Fig. 2A).

As the elevated temperature treatment was only applied post-anthesis, the grain set was unaffected by the high temperature treatment. In contrast, nitrogen input did influence grain setting, with fewer grain being produced when nitrogen supply was limited. This shows how the effect of decreased nitrogen has a twofold impact on grain yield: reducing both

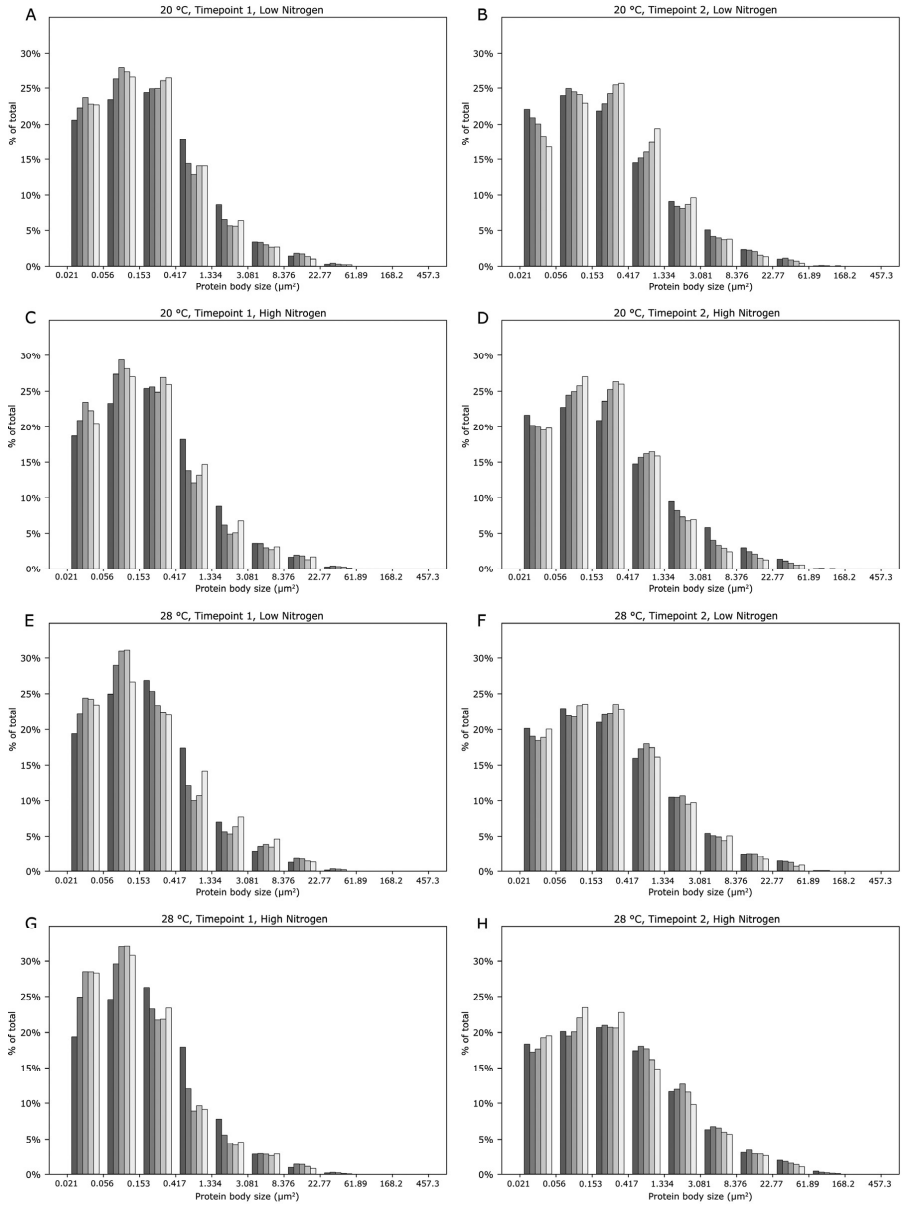


Fig. 5. Histograms showing the frequency of protein body sizes across each of the five endosperm zones. Grain was (A) grown at 20 °C with low-nitrogen treatment and sampled at timepoint 1, and (B) sampled at timepoint 2; (C) grown at 20 °C with high-nitrogen treatment and sampled at timepoint 1, and (D) sampled at timepoint 2; (E) grown at 28 °C with low-nitrogen treatment and sampled at timepoint 1, and (F) sampled at timepoint 2; (G) grown at 28 °C with high-nitrogen and sampled at timepoint 1, and (H) sampled at timepoint 2. Data for zones 1–5 (see Fig. 1) are represented within each graph by shades of grey from dark (zone 1) to light (zone 5), and are the same zones that were used to plot the means in Fig. 4.

grain number and grain size. However, the yield reduction caused by limiting nitrogen input can be primarily accounted for by the reduction in grain set rather than a decrease in grain size.

Grain nitrogen

Elevated post-anthesis temperature resulted in an increase in grain protein concentration, whilst limiting the supply

of nitrogen during vegetative development had the result of decreasing the concentration of protein in the mature grain. The effect of decreased nitrogen supply on the grain protein concentration was greater in plants subjected to high post-anthesis temperatures than those grown under control temperatures, suggesting an increased sensitivity to nitrogen input at higher temperatures, and confirming previous reports (Bhullar and Jenner, 1985; Gooding, 2010). However, elevated post-anthesis temperature treatment did not have an impact on the yield of protein in the grain: the same physical amount of protein accumulated in the grain regardless of temperature. Hence, elevated temperature during grain-filling is primarily impacting on the accumulation of starch, resulting in a greater concentration of protein in the grain, as previously reported by Altenbach *et al.* (2003), Dupont *et al.* (2006), and Koga *et al.* (2015).

Microscopy analysis

The analysis of developing grain taken from the centre of the ear using light microscopy showed that both post-anthesis temperature and nitrogen supply interacted to affect the accumulation of protein within the endosperm during grain-filling, and that this was accompanied by changes in the size distribution of protein bodies.

Higher post-anthesis temperature resulted in grain with higher overall protein concentration and steeper protein gradients across the grain, i.e. protein was disproportionally accumulated in the outer layers of the endosperm. These outer endosperm cells may adhere to the aleurone layer during milling, and hence if heatwave conditions were experienced during grain-filing it is possible that high grain protein content may not result in a proportionally high protein content of white flour. Consequently, depending on the flour extraction rate, grain nitrogen measurements for crops that are exposed to high temperature during grain-filling may have a lower predictive value of flour functionality than for crops grown at lower temperatures. Since limiting nitrogen supply appeared to have the opposite effect, the increase in protein concentration gradient caused by high post-anthesis temperatures could be offset to a certain degree by decreasing the amount of nitrogen applied to the crop prior to anthesis. Although it is not possible to accurately predict weather conditions during grain-filling at the point of applying nitrogen to the crop during vegetative growth, data from this study may prove useful in cases where post-anthesis applications of nitrogen to wheat crops may have been scheduled and, more generally, in the context of years where higher summer temperatures are expected (for example, in coincidence with El Niño/North Atlantic Oscillation weather phenomena). Reduction of nitrogen input, however, may add to the reduction in yield already observed in wheat grown under higher temperatures.

Size distribution analysis of protein bodies in the grain further explained the nature of the observed gradient in protein concentration described in this experiment, and the distribution was characterised by a general trend of decreasing protein body size towards the central endosperm. However, variation in this general trend was observed, with the analysis of the earlier sampling timepoint showing an increase in mean protein body size towards the central endosperm after the initial decrease. Analysis of the histograms in Fig. 5 shows that this was due to an increase in the proportion of medium-sized protein bodies in the inner endosperm, accompanied by a decreased proportion of small protein bodies, with the frequencies of larger protein bodies showing minimal differences across the endosperm zones. It is possible that this effect was caused by the levels of starch in the grain, with lower levels of starch accumulated at the earlier sampling timepoint allowing the protein bodies to expand to a greater size in the central endosperm tissue.

The size distribution analysis revealed that protein bodies were generally larger at the later sampling timepoint, which is to be expected since grain protein production is ongoing during grain-filling, and smaller protein bodies will fuse to form larger protein bodies as they grow (Moore *et al.*, 2016). It was apparent that most of the measured protein bodies were very small, as shown in Fig. 5, with all treatments showing an abundance of small bodies and comparatively few larger bodies. However, it was in the frequency of large protein bodies that the differences were most apparent, with more larger protein bodies observed at the later sampling timepoint, and with distinct gradients in the proportions of the larger protein bodies across the endosperm zones.

Conclusions

Both grain morphology and nitrogen concentration were affected by elevated post-anthesis temperature and restricted nitrogen supply. High post-anthesis temperature resulted in smaller grains with a higher concentration but same overall content of nitrogen in the mature grain, whilst limiting nitrogen supply resulted in both smaller and fewer grains being produced, and decreased both grain nitrogen concentration and nitrogen yield.

Furthermore, temperature and nitrogen supply interacted to affect both total protein distribution and protein body size distribution, as detected through image analysis of microscopy sections from grain at mid grain-filling. Elevated post-anthesis temperature affected both the gradient of total protein within the wheat grain endosperm, and the size distribution of protein bodies, with an increase in the steepness of the negative gradient observed between outer and inner endosperm tissue. Increasing nitrogen supply also had comparable, but less pronounced, effects on the gradient of both total protein and protein body size. Whilst we cannot comment on the changes to baking quality of individual mill streams that these gradients may produce, this study shows that high temperature during grain-filling has the capacity to reduce the protein yield of white flour, as proportionally more protein will be adhered to the aleurone layer, which is removed during the production of white flour. However, we also show that this potentially negative effect may be somewhat reduced by a reduction in the supply of nitrogen fertiliser to the plant prior to grain-filling.

Acknowledgements

Rothamsted Research receives strategic funding from the Biotechnological and Biological Sciences Research Council (BBSRC) as part of the 20:20 Wheat (BBS/E/C/00005202) and Designing Future Wheat (BBS/E/C/000I0220) projects. George Savill is a PhD student with the University of Reading, and is funded by the BBSRC, Lawes Trust, and the University of Reading. The authors wish to thank Peter Shewry for many helpful suggestions.

References

Alexander LV, Zhang X, Peterson TC, *et al.* 2006. Global observed changes in daily climate extremes of temperature and precipitation. Journal of Geophysical Research **111,** D05109.

Altenbach SB. 2012. New insights into the effects of high temperature, drought and post-anthesis fertilizer on wheat grain development. Journal of Cereal Science **56,** 39–50.

Altenbach SB, DuPont FM, Kothari KM, Chan R, Johnson EL, Lieu D. 2003. Temperature, water and fertilizer influence the timing of key events during grain development in a US spring wheat. Journal of Cereal Science **37,** 9–20.

Bhullar SS, Jenner CF. 1985. Differential responses to high temperatures of starch and nitrogen accumulation in the grain of four cultivars of wheat. Australian Journal of Plant Physiology **12,** 363–375.

Cobb NA. 1905. Universal nomenclature of wheat. Sydney, New South Wales: Department of Agriculture, Miscellaneous Publication No. 539.

Dawson CJ, Hilton J. 2011. Fertiliser availability in a resource-limited world: production and recycling of nitrogen and phosphorus. Food Policy **36,** S14–S22.

Dupont FM, Hurkman WJ, Vensel WH, Tanaka C, Kothari KM, Chung OK, Altenbach SB. 2006. Protein accumulation and composition in wheat grains: effects of mineral nutrients and high temperature. European Journal of Agronomy **25,** 96–107.

Fischer RA. 1985. Number of kernels in wheat crops and the influence of solar radiation and temperature. Journal of Agricultural Science **105,** 447–461.

Fisher DB, Macnicol PK. 1986. Amino acid composition along the transport pathway during grain filling in wheat. Plant Physiology **82,** 1019–1023.

Gooding MJ. 2010. The effects of growth environment and agronomy on grain quality. In: Weigley C, Batey I, eds. Cereal grains: assessing and managing quality. Cambridge, UK: Woodhead Publishing Ltd, 393–412.

He J, Penson S, Powers SJ, Hawes C, Shewry PR, Tosi P. 2013. Spatial patterns of gluten protein and polymer distribution in wheat grain. Journal of Agricultural and Food Chemistry **61,** 6207–6215.

Hennessy KJ, Fawcett R, Kirono D, *et al.* 2008. Drought: exceptional circumstances. An assessment of the impact of climate change on the nature and frequency of exceptional climatic events. Melbourne/Victoria: Australian Government Bureau of Meterology/CSIRO.

Kent NL. 1966. Subaleurone endosperm cells of high protein content. Cereal Chemistry **43,** 585–601.

Koga S, Böcker U, Moldestad A, Tosi P, Shewry PR, Mosleth EF, Uhlen AK. 2015. Influence of temperature on the composition and polymerization of gluten proteins during grain filling in spring wheat (*Triticum aestivum* L.). Journal of Cereal Science **65,** 1–8.

Moore KL, Tosi P, Palmer R, Hawkesford MJ, Grovenor CR, Shewry PR. 2016. The dynamics of protein body formation in developing wheat grain. Plant Biotechnology Journal **14,** 1876–1882.

Public Health England. 2014. Heatwave plan for England. London: Public Health England, publication no. 2015049.

Tosi P, Gritsch CS, He J, Shewry PR. 2011. Distribution of gluten proteins in bread wheat (*Triticum aestivum*) grain. Annals of Botany **108,** 23–35.

Ugalde TD, Jenner CF. 1990a. Substrate gradients and regional patterns of dry matter deposition within developing wheat endosperm. I. Carbohydrates. Australian Journal of Plant Physiology **17,** 377–394.

Ugalde TD, Jenner CF. 1990b. Substrate gradients and regional patterns of dry matter deposition within developing wheat endosperm. II. Amino acids and protein. Australian Journal of Plant Physiology **17,** 395–406.

Wan Y, Gritsch CS, Hawkesford MJ, Shewry PR. 2014. Effects of nitrogen nutrition on the synthesis and deposition of the ω-gliadins of wheat. Annals of Botany **113,** 607–615.

Wang YG, Khan K, Hareland G, Nygard G. 2007. Distribution of protein composition in bread wheat flour mill streams and relationship to breadmaking quality. Cereal Chemistry **84,** 271–275.

Wheeler T, von Braun J. 2013. Climate change impacts on global food security. Science **341,** 508–513.

Yahata E, Maruyama-Funatsuki W, Nishio Z, Yamamoto Y, Hanaoka A, Sugiyama H, Tanida M, Saruyama H. 2006. Relationship between the dough quality and content of specific glutenin proteins in wheat mill streams, and its application to making flour suitable for instant Chinese noodles. Bioscience, Biotechnology, and Biochemistry **70,** 788–797.

Zhou Q, Li X, Yang J, Zhou L, Cai J, Wang X, Dai T, Cao W, Jiang D. 2018. Spatial distribution patterns of protein and starch in wheat grain affect baking quality of bread and biscuit. Journal of Cereal Science **79,** 362–369.